

揭示多元資訊環境下的意見氣候分析新途徑*

蔡依霖**

投稿日期：2024 年 1 月 31 日；通過日期：2024 年 7 月 15 日。

* 本文作者在此衷心感謝兩位匿名審查者與編委會對論文提供的寶貴意見。

** 蔡依霖為國立政治大學傳播學院博士，2025 年 2 月起任職於淡江大學大眾傳播學系助理教授，e-mail: elaine019691@gmail.com。

本文引述格式：

蔡依霖 (2025)。〈揭示多元資訊環境下的意見氣候分析新途徑〉，《新聞學研究》，162，97-155。https://doi.org/10.30386/MCR.202501.0003

《摘要》

本文重新解構 Noelle-Neumann 在沉默螺旋理論中的要素「準統計官能」(Quasi-Statistical Organ)，將原本理論所處的媒介環境修正為當今多元的資訊環境，透過平臺的特性與介面展現的集體用戶表現(Aggregate User Representation)發展出更貼近批踢踢實業坊(PTT)使用者「鄉民」視角的資料蒐集途徑，並利用語料庫分析取徑萃取出能夠提取、預測意見氣候的探測語(Probes)，藉此發展出能夠不先預設觀測範圍的意見氣候分析架構。

資料結果顯示，本文還原準統計官能的原理，利用介面的集體用戶表現與修正過後的「探測語」概念設計出可以持續更新的資料蒐集方式，並搭配自建語料庫分析的方式，能建立一個可持續更新探測語的分析架構，且實測以此架構找出的探測語確實能發揮預測意見氣候的效果。本文認為從介面與準統計官能的角度建立的意見氣候分析框架，將更能夠因應瞬息萬變的網路資訊環境，且探索過去預先劃定觀測範圍所不能觀測到的細微現象。

關鍵詞：集體用戶表現、意見氣候、準統計官能、語料庫分析

壹、研究背景與動機

Noelle-Neumann 在 1972 年提出沉默螺旋理論以後，陸續收穫了許多針對理論的批評，Scheufele & Moy (2000) 匯整了沉默螺旋理論與研究設計的矛盾處，統整了學界一直以來對於沉默螺旋理論解釋力不足的聲音。Scheufele & Moy 整理了 Noelle-Neumann 的研究論述後認為，沉默螺旋理論有三大爭議，分別是基本假設不明確、操作型定義不完全以及忽略了更宏觀的變數，其中基本假設裡對「準統計官能」(Quasi-Statistical Organ) 的切入視角就影響了輿情研究的大方向。

Noelle-Neumann 在理論中提到因為害怕被孤立而啟動的準統計官能是指閱聽人具備覺察主流意見的直覺，然而這個直覺源自於閱聽人透過吸收所處環境的資訊拼湊出主觀的主流意見，但後續許多研究卻直接以民調、選舉結果的客觀的數據結果與之混為一談，一方面是因為定義不夠清楚，另一方面是因為研究者難以蒐集閱聽人主觀視角中的主流意見樣貌。在 Noelle-Neumann 提出的沉默螺旋理論中，指出媒體形塑的民意與民眾感受到的民意會有歧異，由於多數人們具有害怕被孤立的社會性，因此在與媒體環境互動時，會評估自己的意見與主流意見的差距，來決定是否發表意見，而這個過程就是意見氣候形成的雛型。根據 Noelle-Neumann 的論點，民意與意見氣候的核心是「閱聽人覺察強勢意見」的過程，然而意見氣候存在於閱聽人覺察後的結果，因此這種意見氣候必須事後對受試者訪談才得以蒐集。時至今日社群媒體蓬勃發展的時代，我們能夠透過蒐集社群媒體的數據觀測意見氣候，也使得不同的行動者演化出了不同的感知方式，Scheufele & Moy 對準統計官能定義與資料蒐集的問題也有了新的解方。閱聽人腦中的意見氣候圖像與研究者

能觀測到的意見氣候樣貌在網際網路上更加靠近。不論是 Facebook 在 2012 年執行的情緒實驗描繪出人們如何透過使用社群媒體接觸到意見氣候 (Kramer et al., 2014)，或是 PTT 的網友也流傳著一句「拔草測風向」，都反映出人們是透過觀察其他人持續製造的意見來推測意見氣候將如何發展，至此，過去研究視角與閱聽人視角不一致的問題開始有機會整合。

蒐集意見氣候的行動者 (例如政府、企業、媒體機構以及研究者) 能夠以更全面且量化的方式觀測，過去必須透過民意調查蒐集的意見氣候，現在則可以利用資料探勘的方式進行輿情觀測與分析，透過觀察社群媒體上撈取到的辭彙來觀察議題的聲量，試圖拼湊出意見氣候的全貌。例如政府對於意見氣候的感知，早在 2010 年就開始對網路輿情進行蒐集，根據研考會於 2010 年製作的委託研究報告《政府重大議題網路輿論趨勢調查研究-以死刑為例》、2011 年針對教育議題製作的委託研究報告《網路輿論意向分析機制之建構與實證研究》，政府感知意見氣候的方式是透過撈取社群網路噗浪 (Plurk) 上的資料對特定議題分析，不過並非廣泛地撈取資料，而是以特定議題的關鍵字蒐集相關的討論進行分析。媒體機構例如新聞媒體，或是資訊媒體如網路溫度計、QSearch、意藍數據等，同樣會蒐集社群媒體上的資訊，以各自的方式計算 (聲量、情緒分析) 後推測出整體的意見氣候。而研究者對於社群意見的感知則是隨著研究方法的發展而有不同，從訪談、網路民族誌到資料探勘，便能夠針對欲研究的主題進行意見氣候的感知。不過這些利用軟體廣泛的、超越人力極限的蒐集方法，與閱聽人的視角仍然無法完全重合，閱聽人無法廣泛地一次瀏覽上百個網路頁面再統計出當今最熱門的議題，也難以觀測出以詳細數據比較分析的結果，因此需要設計出一套更加貼合閱聽人視角的觀測輿情的方法。

過去的輿情觀測方式礙於科技只能以二手的方式蒐集閱聽人所見，當今由於科技突破限制卻又超越了閱聽人注意力所及範圍，從研究者的視角更面臨了研究結果難以概括至其他主題的問題。當今針對意見氣候的觀察與研究，都是透過劃定主題、蒐集資料以後，分析推測出整體的意見氣候。而以網路資料為基礎的蒐集資料方式，大多以關鍵字進行篩選，接著對詞頻、情緒分析、共現等方式觀測意見氣候。這樣的研究取徑所獲得的結果都只能針對單一議題進行詮釋，使得對閱聽人行為的研究結果難以概括至其他相關的議題。

此外，目前的研究都需要等欲觀察的主題已經結束以後才能針對劃定範疇的資料進行詮釋。然而意見氣候是一個動態的過程，將輿論行為包裹在已經結束的事件來觀察，也同樣會造成結果難以概括至其他議題。而且當研究者在蒐集資料時，時常需要閱聽人透過問卷自己回顧面對意見氣候時的媒體使用行為，或者透過撈取資料來歸納閱聽人的反應，但這樣步驟都不是直接觀察「意見氣候形成」的過程，就有可能與實際身處於意見氣候當中的網友所見不相符。

因此，本文將探討下列問題：（1）新媒體環境中準統計官能的樣貌為何；（2）如何依照準統計官能的邏輯蒐集資料；（3）依照準統計官能設計的觀測方式是否具有預測性。

貳、文獻探討

一、媒介環境變遷所致的理論裂縫

沉默螺旋理論詳細的建構時空背景與契機已經有文獻做出很完整的脈絡梳理（翁秀琪，1990；孫秀蕙，1993）。Noelle-Neumann 觀察 1965

年西德大選時，發現民調結果在最後一刻出現剪刀開口，取得強勢民意的政黨勝選，往後的選舉也觀察到同樣的結果，Noelle-Neumann 認為擁有強大聲量的少數會使得沉默多數選擇不表態，而媒體報導會增強這樣的問題，使得沉默螺旋更加嚴重（翁秀琪，1990）。

在當時的媒介環境下，閱聽人除了透過觀察人際傳播的結果以外，只能透過攝取大眾媒體的報導來推斷意見氣候的樣貌（翁秀琪，1990），正如當時西德的選民只能透過觀察民調結果與媒體報導來推斷目前哪個政黨比較得勢，進而決定是否與他人談論、表態甚至投下選票。其中媒體的影響力很大，並且選民除了投票以外，對意見氣候的反應呈現也受限於當時的調查方式，例如只能在訪談中設想是否願意在火車上與同一車廂的陌生旅客表態自己的政治選擇（Noelle-Neumann, 1973, 1984；轉引自翁秀琪，1990），或是參加民意調查。在當時的媒介生態下，使得閱聽人只能接收以民調與媒體報導形塑的意見氣候，並且閱聽人在其中的作用力比較不直接。

然而在沉默螺旋理論提出以後的十年之間，隨即出現針對理論的批評，包括忽略參考團體的重要性、過分強調害怕差異、忽略個體差異、準統計官能的解釋不足（翁秀琪，1990），例如 Glynn & McLeod（1984）就認為並不是所有人都會在發現自己的意見逆風時都會選擇沉默，因此提出了「死硬派」（hardcores）的存在，Salmon & Kline（1983）則認為不一定只有準統計官能會影響閱聽人感知主流意見，也可以從心理學的鏡中之我（looking-glass theory）來感知他人的想法，這時候出現的大多是針對閱聽人最後落入沉默螺旋此結論的批評。Scheufele & Moy（2000）則在千禧年的時空背景下重新審視沉默螺旋理論提出的 25 年以來，歸納出的三大爭議，分別是基本假設不明確、操作型定義不完全以及忽略了更宏觀的變數。Scheufele & Moy 認為沉默螺

旋理論在不同的環境之下似乎不完全能適用，例如提出沉默螺旋時德國的時空背景、民族性等變項都與美國不同，此外所獲取的研究數據（survey data）也需要根據所在的環境試著蒐集更宏觀的資料。Noelle-Neumann 的門生 Donsbach 以及另外兩位學者 Salmon 與 Tsfaty 則編纂了論文合輯 *The spiral of silence: New perspectives on communication and public opinion*，特地從方法論、網路時空背景以及相關的研究重新審視沉默螺旋理論（Donsbach et al., 2014）。網路時空背景的討論，則著重於網際網路環境讓閱聽人發表意見的門檻降低，因此閱聽人能夠在匿名環境下發表意見，不需要直接面對被孤立的恐懼，然而這些討論仍然著重於閱聽人在偵測完主流意見的樣貌以後是否會落入沉默螺旋。

當今的意見氣候形成已經不像過去閱聽人只能等待主流媒體與民調「製作」好意見氣候的樣貌，而是可以用非常方便且趨近於匿名的方式任意地表達意見，任何立場的閱聽人都有機會成為「強勢意見」，而且發現自己與強勢意見相左的閱聽人也不必然會擔心與眾不同而選擇沉默。於此，啟動準統計官能之後，閱聽人的選擇就與過去的理论出現差異了。由於電腦中介傳播（Computer-mediated Communication）具有視覺匿名（visual anonymity）與管道有限（limited channel）的特性，人們在網路上更容易揭露自我（self-disclosure）（Joinson, 2001），因此過去驅動沉默螺旋理論的其中一個強烈的動力「害怕被孤立」也就得以解套。甚至在網路上有些人比在現實中更頻繁或更強烈地進行自我披露或行為表現，Suler（2004）將這樣的現象稱為網路去抑制效應（online disinhibition effect），因為網際網路上分離性匿名（dissociative anonymity）、隱形性（invisibility）、非同步性（asynchronicity）、唯我論內化（solipsistic introjection）、分離性想像（dissociative imagination）和權威的最小化（minimization of authority）這六大特性，

人們更願意吐露心聲。

當線上的發言在電腦中介傳播的諸多特性之下讓意見的呈現能夠與線下的自我切割，也就不需要擔心表態意見會直接地導致被孤立。因此在啟動準統計官能以後，閱聽人的行為選擇就變得多元了，人們可以主動地群聚討論、分享資訊、反饋意見，也可以安靜、被動地增加集體用戶表現（aggregate user representations）（Walther & Jang, 2012），例如替瀏覽人次增加一個數量，或是成為眾多同時上線人數的一員，在意見氣候生成的過程中閱聽人獲得了更多的主動性。

更甚者，當今有許多不同特性的社群平臺，閱聽人可以選擇論壇式的平臺面向議題圍觀，也可以選擇在以自我為中心的平臺發表意見。當閱聽人選擇面向議題圍觀時，能夠發展出更多不同的行為，例如在議題之間形成流動的群聚（黃厚銘、林意仁，2013），或者浸淫在與自己意見相似的資訊環境形成迴聲室（echo-chamber, Sunstein, 2009）、同溫層（Flaxman et al., 2016），進而與其他聲音分道揚鑣、逐漸極化（Lee, 2015），因此，面對意見氣候時閱聽人不再只能設想自己是否願意和同一節車廂中的其他陌生乘客討論當今熱議的話題，所做出的行為也不再只能選擇是否順應沉默螺旋，這也讓沉默螺旋的結論有了新的出路。

二、新媒體環境下的社交線索

過去不論是輿論或者意見氣候，都是包裹在「民意」當中的概念。國內學者曾虛白於 1969 曾於〈民意測量質量之辯〉一文中彙整民意的定義，包括民意即「意見的氣候」以及「意見的發展程序」，前者是為了凸顯民意具有瞬息萬變的特質，後者則強調了民意是許多事件與調查結果不斷生成且彙整的過程，並將此過程以氣態、液態、固態的「結晶

性」劃分階段，指稱民意會從眾說紛紜的氣態階段，經過不同立場分化成液態階段，直至最後經過各種大事件（諸如形成公約、政治人物發表演說、公眾人物的公開辯論等）的定調成為固態階段。

從曾虛白的論點中可以理解到意見氣候是動態的，與後續西方學者 Noelle-Neumann 所提出沉默螺旋不謀而合，在 Noelle-Neumann 的研究中，意見氣候指的是環境中意見分布的狀態，而此狀態又是一動態過程「沉默螺旋」的一部分。延續著沉默螺旋的理論方向，國內學者也持續地以實證研究體現沉默螺旋的過程，例如以興建核四廠議題來驗證與沉默螺旋相反的死硬派投射效果（孫秀蕙，1994）或是以臺灣的臺北縣長選舉檢驗沉默螺旋理論（翁秀琪，1997），這些研究也都有提及閱聽人會利用大眾媒體觀察意見氣候，以及觀察意見氣候形成的過程會使得閱聽人做出什麼決定，最後落入沉默螺旋或是成為死硬派。因此我們可以將輿論理解成經過人為匯集成定論的公眾意見（例如寫成報導、作成調查）將意見氣候視為意見正逐漸匯集成「輿論」的動態過程，而準統計官能則是閱聽人偵測社交線索的能力。

（一）「鄉民」眼中的社交線索

在過去的文獻中，閱聽人眼中的意見氣候樣貌大多需要由受訪者回憶自己看過哪些發言、推測出什麼樣的意見氣候，然而在網路環境下研究者也可以一窺多數閱聽人所見的意見氣候。例如觀察 PTT 上 NBA 板的網路群聚起鬨的行為時，暴漲的看板人數、迅速增長的回覆數量便是在網路上集體亢奮的表現（黃厚銘、林意仁，2013）。此外，由於社群媒體能夠在重大議題發生時迅速地聚集大量群眾，在觀測 Twitter 上與選舉議題參與討論的情形時，「留言暴增」在研究中被視為重要的研究發現（鄭宇君、陳百齡，2014）。累積速度快的功能，是能夠更容易地

被其他閱聽人看見，當文章下方的留言以很快的速度累積，將有助於文章受到所屬平臺上社群的關注（李柏彥等，2021 年 6 月 25-27 日）。

而使用者生產內容（UGC）的數量累積與上述的累積速度一樣會讓訊息更容易受到關注，在 Noelle-Neumann 的沉默螺旋理論中，提到閱聽人會利用準統計官能接收大眾媒體與二級傳播的資訊來推估整體的意見氣候，然而在 Web 2.0 的媒體環境中，閱聽人除了直接接觸到其他閱聽人所產製的資訊外，也有更多的社交線索（Social Cues）可以參照，藉此判斷出更具體的意見氣候（Lee et al., 2022）。然而關於更具體的探究在網際網路上的意見氣候與社交線索為何，又會發現不同的場域所能夠呈現出的內容也大相逕庭。Walther & Jang（2012）彙整出社群媒體上呈現廣義整體意見氣候的樣式，大致可以分成三種主要的類別，分別是原著內容（proprietor content，以下簡稱 PC）、集體用戶表現（aggregate user representations，以下簡稱 AUR）以及使用者生產內容（User-generated content，以下簡稱 UGC）。

原著內容指的是例如網路新聞的報導本身，或者是像 Facebook 這種社群媒體中一則一則的貼文；集體用戶表現（AUR）則是指以數量或系統方式呈現出集體意志的訊息，可以分成有意生成（deliberate）的與附帶生成的（incidental），前者指的是例如 Facebook 提供按讚、留言、分享，後者則是指集體在社群媒體上的行為被統計出的數字，例如 YouTube 影片下方的觀看人次；最後使用者產生內容（UGC）指的則是例如對應於網路新聞，下方的網友留言，或是社群媒體上針對原著內容下方的評論。針對這三種不同的內容樣式，Neubaum & Krämer（2017）指出人們在感知意見氣候時，最在意的社交線索（Opinion Cues）正是 AUR 與 UGC，雖然並不能觀測出「意見氣候與自己是否相左」與「願不願意表達自己的意見」之間有顯著的因果關係，卻可以觀

察到人們在回想自己所看到的社交線索時，記得最清楚的總是 AUR 與 UGC。後續 Leong & Ho (2021) 的研究延伸了 AUR 與 UGC 的概念，發現如果 AUR 與 UGC 一起出現，閱聽人更有辦法拼湊出意見氣候整體的樣貌，而其中該研究中的 AUR 指的是 Facebook 上的讚、愛心、生氣、加油與哭臉，而 UGC 則是留言。

從上述的文獻可以理解到在閱聽人眼中的資訊環境裡，當有其他閱聽人的意見迅速、大量地浮現，代表著意見氣候正在形成，而這些大量且迅速浮現的資訊提供了閱聽人足夠的社交線索，使得閱聽人的準統計官能被觸發且開始理解、形塑意見氣候可能的樣貌。

(二) 新聞網站、社交平臺的社交線索

進入網際網路時代以後提到感知意見氣候的研究領域非常廣泛，從口碑研究 (Electronic Word-of-Mouth; eWOM) 領域到網友行為研究 (公民參與、社會運動、選舉研究等)，民眾感知意見氣候的來源比以往更加多元，不只是以往單純的新聞報導這種形式的資訊，而是 PC、AUR 與 UGC 混合的資訊，例如有網友留言的新聞網站、包含瀏覽次數與評論的影音網站、可以看到按讚數的社群媒體等，不同的網站介面也會啟發不同的研究重點。

早期有一篇最接近 AUR 的研究，是 Knobloch-Westerwick 等人研究新聞入口網站的新聞評分機制是否影響閱聽人選擇性暴露在特定的報導下 (Knobloch-Westerwick et al., 2005)，不過此時的新聞網站還沒有 UGC 的功能。在 eWOM 領域，其他消費者的評價對於閱聽人的說服效果就有很大的影響，Dai et al. (2019) 的研究就分別探索主流意見氣候與少數意見對於瀏覽者的影響力為何，研究的對象為飯店評論網站 Tripadvisor 的 PC (網友對飯店的評價) 與 AUR (飯店總體獲得的星

級），由於介面會特別強調其他網友的推薦星級與留言，因此不論是主流意見（五顆星）或是少數意見（一顆星）都會非常顯著，因此有些時候負評的影響力也非常顯著。Dai 等人的研究將「閱聽人面對不確定性會尋求資訊」的過程具象化了，人們透過 PC 與 AUR 來降低不確定性，而且影響人們抉擇的 PC 與 AUR 是可以觀測的。

關於 Facebook 介面上的 PC、AUR 與 UGC 也有幾篇相關的研究，Zerback & Fawzi 就特別針對社群網站的形式（online exemplars）是否觸發沉默螺旋進行實驗，研究特別讓受試者回憶是否記得實驗設計中的 Facebook 貼文中，特定文章的按讚數與留言數多寡以及是否願意發布與該文章立場相左的文章，研究顯示受試者對於 AUR 表現突出的文章都會比較有印象（Zerback & Fawzi, 2017），這顯示了 PC、AUR 與 UGC 對閱聽人來說確實是重要的社交線索來源。Neubaum & Krämer 則是創立一個 Facebook 社團讓受試者在社團中發文並討論安樂死合法化與同性婚姻，社團中實驗者有準備了幾則包含 PC、AUR 與 UGC 的文章作為刺激物（Stimulus Material），讓受試者可以看到貼文、hashtag、按讚數與留言，研究發現 AUR 與 UGC 都是幫助受試者構築輿論樣貌的重要社交線索，可見 PC、AUR 與 UGC 是影響民眾感知意見氣候的關鍵（Neubaum & Krämer, 2017）。AUR 提供閱聽人整體意見氣候的輪廓，PC 則提供了個別意見氣候的詳細內容，UGC 則提供了大眾對意見氣候的回饋，因此透過接收 PC、AUR 與 UGC 可以幫助閱聽人構築意見氣候的樣貌與可能的發展趨勢，而研究者也可以觀測到與閱聽人視角相同的數據結果。其中，由於 AUR 與 UGC 是閱聽人花最少的心力可以獲得最廣泛的社交線索，在 Walther & Jang、Neubaum & Krämer 的研究中都分別提到了 AUR 與 UGC 的影響力又更大於 PC。

（三）PTT 介面提供的社交線索

批踢踢實業坊（簡稱 PTT）做為一個綜合各路主題資訊的集散地，性質與 Facebook 這類會被稱之為社交網路（SNS）的平臺不同，Facebook 與 Twitter 這兩個全球主要社交媒體科技公司最初在設計平臺技術結構時，即以促進人們社交互動與訊息流動為主要目的（鄭宇君，2017），而 PTT 作為論壇，則是提供人們談論各種主題的公共空間（黃厚銘等人，2016）。

根據 PTT 官方網站的說明，其本質上是提供讀者查詢各項資料的電子公布欄（BBS, Bulletin Board System）（見 <https://www.ptt.cc/index.bbs.html>），因此整體的設計都是方便使用者主動地瀏覽、查詢自己感興趣的內容，這也使得 PTT 與其他社群媒體有了極大的不同。在 PTT 上是使用者往資訊移動，社群媒體則是靠平臺的演算法將資訊往使用者推送，因此使用者在社群媒體上就無法全觀地瀏覽資訊，而是只能片段地看到平臺依照演算法篩選過後的資訊。

例如 Facebook，即便具有 PC、AUR 與 UGC，但因為其「好友清單」的功能使得閱聽人傳散資訊時永遠「面向他人」，使資訊的傳遞無方向性且有範圍（劉慧雯，2017），PTT 的介面更像是一個通往各地的車站，以話題為月台讓人們得以聚集，這使得觀察意見氣候生成更加直接，不會受到好友清單這類隱私設定、不透明演算法的干擾。

國外與 PTT 比較接近的論壇 Reddits 也有一些相關的研究，同樣包含從機緣的角度與 AUR、UGC 角度的討論。Reddits 上同樣有 PC（貼文）與 UGC（貼文與下方的回覆），此外針對不論是 PC 或 UGC，都可以進行上箭頭（upvote）或下箭頭（downvote）的投票來表達喜歡或不喜歡，因此在每一則貼文與回覆之前，都可以看到以上下箭頭數量為標示的 AUR，而這個 AUR 又具有影響能見度的功能，因此對內容投票的

功能就衍生出一些能供性，包含對內容進行審議、形塑討論的走向（Leavitt & Robinson, 2017），而且 Reddits 上的 AUR 對後來持續生成的 AUR 具有影響力，當實驗者對於先發布的留言投票操作成贊成票或反對票居多，後續的投票結果會與一開始顯示的 AUR 有顯著的一致性（Weninger et al., 2015），這顯示在論壇形式的網站上，AUR 確實會影響後續的意見氣候生成。Davis & Graham（2021）的研究就從能供性的視角，觀察到贊成與否的投票功能給了使用者表達情緒與參與內容的能供性，並發現 Reddits 上的投票機制對於情感正負向的表達有直接的影響。另一篇針對 Reddits 審議式民主功能的研究則發現該 AUR 有時會在討論的過程中發揚或壓制某一個立場的聲音（Buozis, 2019），這兩篇研究都可以看到在這種論壇類型的媒介上，介面的社交線索是影響後續立場表達的重要因素，而 AUR 與 UGC 就是社交線索的展現。

PTT 這個論壇本身的特性也符合 AUR 與 UGC 具備能夠引導意見氣候感知的能供性，有許多社交線索提供給 PTT 的使用者「鄉民」，研究者黃厚銘等人（2016）在《婉君你好嗎？給覺醒鄉民的 PTT 進化史》一書中提到：

除了消息傳得夠快，論壇開放任何人參與的特性，也有助於確保達到足夠「數量」乃至「多元性」的群眾基數，於是，人們便能針對同一個主題，閱讀到大量使用者的看法和意見（p. 50）。

由上述段落可見鄉民的確浸淫在大量意見之中，除了直接閱讀意見，更可以觀測到足夠「數量」的資訊，實際上介面提供哪些 AUR，則必須從登入畫面以及不同的載具來觀測。

在鄉民進入 PTT 的第一個畫面，看到的會是類似入口網站

(portal) 的畫面，這裡第一個出現的數字就是登入當下的線上人數，接著進入分類看板以後，就是第二層的入口網站頁面，接著第三層畫面才能看到更多接近 AUR 的資訊，包括不同主題的板依照人氣的排序、右上方人氣的顏色等；而手機版 PTT 除了同樣也有「熱門看板」的排序，還有「最新消息」與「延燒話題」，這些介面的設計都是「聚集導向」的，不像是 Facebook 或 Twitter 的設計比較偏向鼓勵分享、產製內容與社交，PTT 的介面更加強調「哪裡人多」，且以各種指示（諸如上線人數、熱門排序、熱門程度等），到這個階段為止鄉民接收到的 AUR，只能夠判讀出哪裡可能有意見氣候，但尚未能判斷意見氣候為何。

而各個看板點進去以後，鄉民可以接觸到以數字呈現的 AUR 分別為右上方該板的人氣以及各篇日期前方的推噓文相減後的數量，當推文數量減去噓文數量後大於 100 者，就會在最前方顯示「爆！」，反之相減後小於 -100 者則會顯示「XX」，這樣的資訊除了顯示出意見的聚集以外，就開始初步地顯現出意見的正負傾向，然而推噓不必然等於正面或負面，因此鄉民還需要直接閱讀推噓文的內容才能加以判斷。因此，如果以 Walther & Jang (2012) 的分類方式，PTT 文章所能提供的社交線索分別是：作為 PC 的原始文章、作為 UGC 的原始文章及下方的推噓文內容、以及推噓文數量標示、熱門排行等 AUR。

三、新媒體意見氣候與語料庫方法結合的新研究取徑

前述幾個不同類型的媒體，都有 PC、AUR 與 UGC 的特性，而反映出的都是使用者對於意見氣候反饋的行為。當閱聽人在面對輿論時啟動準統計官能所面對的環境，與研究者觀察輿論時面對的環境在介面上相

會，比起過去只能間接地獲得閱聽人準統計官能的觀察（透過問卷提問或是請受試者自己紀錄），利用介面的邏輯進行撈取資料的研究設計，就可以讓研究者的角度與閱聽人的視角更加貼近。

在前言提到需要蒐集意見氣候的行動者（例如政府、企業、媒體機構以及研究者）在蒐集資料時都是透過從上而下（**top-down**）的取徑在蒐集資料與分析資料，本文將嘗試以從下而上（**bottom-up**）的研究取徑，沿著使用者的視角蒐集資料與分析。從上而下與從下而上這兩種研究取徑曾經被應用在許多不同的學科。例如在公共行政的領域講的是政府治理的方式（朱景鵬、朱鎮明，2006），在傳播領域說的則是傳播的方式（Wilson & Irvine, 2013），甚至語言學的研究中也有以這兩種取徑作為分析比喻的研究方式（Chung et al., 2005）。與本文相似的應用方式出現在資料科學領域中（Dodge, 2021）。以 Dodge 的研究為例，提到在處理人們製造出的龐大數據時，可以用下行的方式處理數據，即以現有對人們行為（**behavior**）的理論把蒐集到的零碎行動（**movement**）進行歸納，或者用上行的方式透過了解所有的行動進而彙整成可預測或可歸納的長期行為。而本文的資料蒐集與分析曲徑除了同樣也是資料驅動（**data-driven**），且試圖想透過瞭解行動並彙整成行為模式，更想讓研究者的視角與閱聽人更加貼近。

若檢視過去對於網友行為研究的「研究者視角」，大多以從上而下的研究取徑來進行論述，其研究步驟的共通點即「個案」或「主題」先行。例如學者鄭宇君與陳百齡以 2012 年總統大選為關鍵個案，劃定主題之後決定以政黨及參選人作為關鍵字撈取資料（鄭宇君、陳百齡，2014）。這樣的方法可以更針對研究者想討論的浮現社群進行觀察，然而對於身在輿論中的使用者，並不一定會使用設定好的詞組討論該關鍵個案。另一案例為學者傅文成與陶聖屏以 2013 年洪仲丘事件為案例，

設定七個與洪仲丘事件相關的指標議題並列出關鍵字來爬文，透過時間序列與分享次數來呈現謠言的「網路模因」傳播模式（傅文成、陶聖屏，2018）。這樣的方法確實幫助研究者獲得網友傳播訊息的行為紀錄，但事先框定指標議題就排除了其他可能的網友行為，例如若要觀察政治人物柯文哲，研究者若事先設定了「柯文哲」為關鍵字，便不會撈取到後來逐漸發展出的「柯 P」、「蚵 P」等新發明的詞彙。

由上而下的研究視角可以更聚焦於研究者想觀察的主題，不過本文想嘗試的是「非事先指導式」的研究步驟，因為身在輿論中的網友並不會有「個案」觀念，如果要觀察意見氣候的形成過程，並不適合事先框定某個主題。採取由下而上的研究取徑，透過不設定主題所蒐集到的資料往上建構資料呈現的現象會更適合。因此本文試圖將蒐集資料的視角與實際在意見氣候中的使用者視角重疊，透過與使用者所見相同的介面，設計資料蒐集的邏輯，進而建立一個更貼近使用者的分析架構。也因為本文沿著使用者的視角蒐集資料，因此除了數字類型的資料（例如文章篇數與留言的數量）也會利用語料庫分析取徑蒐集文字類型的資料。閱聽人在使用 PTT 時，並不會如同研究者能觀測到精確的統計結果，例如包含特定關鍵字的文章總共有多少篇，更多時候是看到相同的主題時常重複出現，或者留言大家常常都在回覆什麼類似的內容，因此出現在留言的詞彙也是本文觀測 AUR 重要的一部分。

關於網路輿論的各種現象與議題過去也有若干研究且主題非常多元，諸如網路新聞學、政治網路口碑、網路霸凌、虛擬社群、同志議題等，這些研究以各種不同的方式呈現出各自關懷的現象以及背後的意義，然而這些以網路論壇為觀察場域的研究都有一個共通點，即其研究結果無法在不同的個案、議題、主題間得到相互驗證，彙整其無法挪用的原因可歸納為以下三點：

(一) 觀察的能供性使得研究結果無法概推

以網路論壇作為觀察場域的研究大多建立於研究者對特定議題的關懷，例如性別、疾病、政治等。鄭芳婷（2016）以類似田野觀察的方式深度解析 PTT 中「拉板」、「彩虹山莊」與「Queerology」幾篇具代表意義的網友發文，然而在其文末特別註明由於議題的特性，使得研究內容並不適合用來代表整個群體，否則會形成論述的霸權。徐士傑等（2013）則觀察 PTT 中「抗癌板」中網友發文所顯現出的社群支持，透過在抗癌板發送問卷了解網路論壇以何種方式提供病友社會支持，然而該篇文章也提到受限於研究方式，其結果僅能代表有效問卷所記錄時期之結果，不能概推到其他時期或其他場域。林思平（2017）則透過深度訪談分析 PTT 八卦板的使用者如何理解並建構網路資訊與網友論述中的真理，然而研究者也註明，受限於研究方法因此研究結果需要更進一步擴充才能適用到其他人口學背景的群體上，此外以遊戲語言的研究角度也可能觀察到與用公共領域角度迥異的結論。

然而主題無法概推的原因，從本文前述社群媒體能供性的觀點，認為是因為不同主題研究關注的能供性不同。例如鄭芳婷（2016）在文中指出 PTT 中的拉板與彩虹山莊在 LGBTQ 族群的使用下，體現了數位離散（digital diaspora）以及「浸淫幻覺」（immersivity）的數位能供性，進而影響了女同志族群的認同過程。徐士傑等（2013）的研究則指出 PTT 中的抗癌板在板友的使用下，發揮了「促進行動型」與「滋養型」社會支持功能。林思平（2017）的研究則提到 PTT 八卦板的板友們在參照資訊、相互溝通，並利用推噓文對文章鼓勵或懲處等行為，構成了真理政治的社會權力實踐。以上述三篇文獻來看，同樣都是 PTT 電子布告欄系統，在不同的研究中卻因為板友對不同主題的看板實踐出相異的能供性。這說明了不同主題所觀察到閱聽人對 PTT 填充的整全補充物

(complementary properties) 不同，因此導致不同主題的研究結果無法概推，探討平臺提供 LGBTQ 在論壇上意識型態的論述功能、抗癌板提供的情感支持功能與八卦板提供網友建構論述的功能不能一概而論。本文認為應該觀察更後設的平臺能供性，在各個不同主題看板由於介面相同，在閱聽人將看板填上各式各樣的整全補充物之前，都會先就著介面實踐出「探索社交線索」的能供性，因此從介面為起點才能找到不會受限於主題的輿論分析框架。

(二) 受限於研究設計的特性

上段提到的限制都不是以語料庫作為研究取徑的文章，若聚焦於蒐集、分析語料的研究，仍然可以觀察到結果無法概推的問題。郭文平（2015）透過建立語料庫後針對新聞報導中提及的經濟概念進行分析，紀慧君（2018）同樣透過建立語料庫分析新聞報導中對於有機食物的再現，趙玉娟與陶振超（2018）則是透過既有的正負向辭典對 PTT 上與政治人物相關的網友貼文進行情感分析，這三篇針對不同主題的研究不約而同地提到，由於必須針對論文題目提出考察與分析，因此不論是蒐集文本時研究者所選擇的關鍵字，或是自建辭典時研究者所構築的語料庫都得依照主題、時空環境調整才能夠扣緊該主題的核心關懷。這樣的研究方法確實會對主題有更細緻的觀察與分析，但共同點即為必須等事件結束之後才能進行研究，否則可能會因為觀察主題發生變化而造成分析被推翻。然而，如果是從介面的 PC、AUR、UGC 切入，由於各個看板的介面並無不同，也不會隨時變化，在觀察時就不會落入某一個特定族群或主題在 PTT 上實踐的行為，也不會造成整全補充物無法概推的情況，而更有機會找到能夠跨越主題的分析方式。

(三) 跨越主題的輿論行為觀察？

不論是否使用語料庫分析方法對文本進行分析，只要是選定主題後再分析的研究通常都無法把結果在其他領域順利重現，正如同上述研究中在不同主題之下的網友行為會因為選擇的主題不同，而使得 PTT 上的同志群體網友、病友團體網友、八卦板網友、政黑板網友所呈現出的行為毫無相似之處。鄭宇君（2017）曾經論述過社交媒體事件之浮現邏輯，研究撈取推特上關於選舉（2012 總統大選）、災難（2015 南港空難）以及社會運動（2014 太陽花運動）的發文與轉推行為，來分析成為社交媒體事件的條件為何，不過文中的論點有對於事件加上一些價值的判斷，例如需要具備問題性、需要影響到線下生活等，在文末也表示因為事先定義了「問題性」，以至於事件之間（例如洪仲丘事件與抵制林鳳營事件）無法類比，然而本研究關心的「網路輿論」並不以是否影響線下生活作為篩選條件，也不是在特定的網路空間（例如某個論壇、或某主題的討論板）才會發生的現象，而是網友在使用網路時的集體行為，因此需要一個能夠不受限於主題的觀察方法才能夠對於網路輿論的行為或現象形成更貼切的觀察與解釋力。

綜觀上述文獻回顧，我們可以瞭解到因為新媒體環境與過去沉默螺旋理論提出時大不相同，理論中提到的準統計官能也有機會做出更多的觀測與應用，本文將結合準統計功能的原理以及語料庫分析取徑，嘗試提出一個更貼近鄉民視角的意見氣候分析架構。

參、研究方法

本文以批踢踢實業坊（PTT.cc）八卦板作為資料來源並依據介面的視覺呈現邏輯蒐集資料。選擇批踢踢實業坊（PTT.cc）八卦板作為觀察

意見氣候領域的理由有三，分別是足夠熱門、不受演算法干擾以及介面本身的特性。

在研究背景就有提到，以流量來看 PTT 是臺灣流量排名前五的社群媒體網路，其中比較容易被主流新聞媒體傳散的社群媒體大多為 PTT 與 Facebook。過去的研究顯示，在新聞來源引用的網路意見中，有標示來源的最大部分即為 BBS 的論壇內容（楊意菁，2013）。其中 PTT 最大的熱門看板「八卦板」更是意見匯聚的重要場域，過去有許多著名的重大政治社會事件皆引起大量網友匯集於八卦板，例如近十年內就曾因為 2013 年的洪仲丘事件、2014 年的太陽花學運、2015 年的八仙樂園塵爆、2016 年的華航罷工等，使得介面出現「綠爆」、「藍爆」甚至「紫爆」的標示（林思平，2017）。即便 2018 年蘇啟誠事件以後，曾經面臨長達兩年半停止註冊，然而截取撰文前日統計，PTT 單日總共上站人次仍然有 6,589,912 人次，平均每個小時就有 274,579 人次在使用 PTT，仍然是意見匯集的重要集散地。

此外，PTT 並不會受到演算法的影響。Meta 旗下的 Facebook 自從 2018 年的劍橋事件（Cambridge Analytica）發生以後，¹ 日趨限縮 API 的取得，即便於 2022 年 Meta 宣布啟動新的數據共享企劃，仍然有許多資訊是不公開給研究者使用的（Ingram, 2022, May 26）。PTT 的程式碼則是公開的，也沒有設置為了推播廣告而影響使用者接收資訊的演算法，因此從研究者的角度可以排除不可控的演算法問題。

最後，PTT 的介面本身具有引導使用者趨向熱門議題的特性。在登

¹ 英國諮詢公司「劍橋分析公司」在未經社群媒體「Facebook」用戶同意的情況下獲取數百萬用戶的個人數據，並利用這些數據操控選舉民意。Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*, 17(1), 22.

入後使用者首先看到的畫面只有 (A) nounce【精華公佈欄】、(F) avorite【我的最愛】、(C) lass【分組討論區】、(M) ail【私人信件區】、(T) alk【休閒聊天區】、(U) ser【個人設定區】、(X) yz【系統資訊區】、(P) lay【娛樂與休閒】、(N) amelist【編輯別名單】與 (G) oodbye【離開，再見】，其中只有 (C) lass【分組討論區】會進入看板列表。進入分類看板以後，會先看到 PTT 建立之初幫使用這初步分類像是入口網站的畫面，分成「市民廣場」、「臺灣大學」、「政治大學」、「青蘋果樹」、「活動中心」、「視聽劇場」、「戰略高手」、「卡漫夢工廠」、「生活娛樂館」、「國家研究院」、「國家體育場」與「--> 即時熱門看板 <--」幾個大項目，若不是有特定要查找瀏覽某個主題的內容，通常使用者會點選最後一個「--> 即時熱門看板 <--」瀏覽當下人數聚集最多的幾個熱門看板。

本文選擇常駐即時熱門看板之首的八卦板作為觀察場域，一方面是其人數上的優勢讓其中的討論足以溢散到其他媒體，另一方面則是八卦板的排序大部分時間都在第一位，在介面上是最容易被使用者接觸到的看板；此外在瀏覽文章時，所有的文章依照時間排序，然而文章數量非常龐大，通常使用者不會逐篇閱讀，介面除了提供關鍵字查詢的檢索方式以外，還會在每一篇文章前顯示推噓文的總和，在本研究中將其視為重要的集體用戶表現 (AUR)。如同觀看 YouTube 時顯示的瀏覽人次，PTT 文章前所顯示的推噓文總合不論是 1 到 99 不等的數字、不同顏色的「爆」甚至噓文過多時會出現的「X」或「XX」都是 AUR 提供給鄉民的社交線索，因此具備這樣的介面使得 PTT 作為觀察「鄉民眼中的輿論」是最適合的。另外，PTT 的文章格式非常固定，包括標題必須分類、必須達到字數下限、轉載新聞必須標注來源等基本規則，因此蒐集到的資料樣貌非常一致，也不需要再進行人工標記或分類。

一、資料蒐集與篩選標準

根據文獻中提到的，形成氣候的意見需要有相對多數的量，以及持續一段時間，此外資訊在不同平臺之間流動亦代表著資訊關注度的提升。過去關於新聞資訊跨平臺流動的研究主要分散於新聞平臺與社群平臺，分別是網友產製的內容往新聞平臺流動的研究（楊意菁，2013；Lee et al., 2005）以及新聞報導往社群平臺轉載的研究（劉慧雯，2020），前者的討論重點在於新聞報導以 UGC 內容作為消息來源是否會把網友的言論公共化，後者則探討新聞小編如何透過轉載新聞與社群對話。新聞機構的社群小編轉載報導的行為除了顯示其個人對於公共話語的重構，轉載的個人也會在介面的引導下，對其所轉載的資訊加註一定程度的意見（劉慧雯，2020），除了新聞機構的小編能這麼轉載新聞，在社群平臺上的「鄉民」同樣也能這麼做。過去關於「分享」資訊的研究指出，當平臺能夠轉載資訊時，便使得閱聽人成為「始終與他人共在」的倫理閱聽人（ethical audience）（劉慧雯，2017），這種考慮其他人對於事件看法進而分享資訊的行為，與本文觀測當代閱聽人透過準統計官能觀測意見氣候的行為不謀而合，因此本文將以此作為資訊蒐集的篩選標準，即蒐集被轉載過且高度討論的新聞，一方面是獲得被鄉民認定其他人也需要看到的、具有公共性的資訊，另一方面也確保該資訊有累積一定的意見數量。

因此本研究利用 PTT-Finder（修改自廖春閔，2018；開發與呈現參見：李柏彥等，2021年6月25-27日）蒐集2022年1月到6月，在PTT的八卦板中，同時具備（1）推文減噓文大於100且（2）標題包含「Re: [新聞]」的文章以及下方的推噓文。推文減噓文大於100的原因是PTT

的介面，當推文噓文大於 100 的話，系統會在文章前方顯示紅色的「爆！」，根據不同的 APP 版本，都會有熱門文章的推送功能，被推爆的文章會優先被使用者看見；² 而選擇「Re: [新聞]」的標題文章，是因為當使用者要回覆一篇文章時需要按下「回覆這篇文章」，此時標題就會有「Re:」，如果是轉貼新聞則需要標注「[新聞]」否則依照板規會被水桶，³ 所以當「Re: [新聞]」一起出現時，代表的是 PTT 上網友認為需要注意的公共資訊獲得持續討論的話題，故作為篩選條件。

蒐集到的資料包含 651 篇文章，319,876 則推噓文，這些資料將會使用中文獨立語料庫分析工具「庫博 CORPRO」進行整理與分析。

表 1：熱門文章資料蒐集則數綜覽

2022 年月份	主文	推文
一月	95	46,915
二月	77	37,706
三月	82	37,035
四月	82	39,853
五月	181	90,421
六月	134	67,946
合計	651	319,876

資料來源：本研究彙整

² 推爆指的並不是推文累積至 100 則，而是將推計算為「+1」、噓計算為「-1」、箭頭計算為「0」，當下方推文持續累計到值超過 100 時稱為「推爆」，被推爆的文章會在標題前方標示紅色的「爆」，故更容易被其他進版的使用者注意到。

³ 水桶指的是禁止發言。

二、利用語料庫語用學（Corpus Pragmatics）設定探測語（probes）

語用學是符號學的其中一個分支。美國哲學家 Charles William Morris 在 1938 年出版的 *Foundations of the theory of signs* 將符號學區分成三大面向，分別是語構學（syntactics）、語意學（semantics）、語用學（pragmatics）。⁴ 簡單地理解，語構學探討的是文法結構中詞彙之間的關係，語意學則是探討詞彙的意義，而語用學則是探討使用者在不同的脈絡下把語言「使用」成什麼意思。如同許多其他人文社會學科，語言學也經歷過實證轉向或經驗轉向（empirical turn, Taavitsainen & Jucker, 2015），使得語用學的研究重要性開始提升。

Taavitsainen & Jucker 綜覽 1990 年以後二十年間，語言學經歷了七大重要的轉向，分別是向語用轉（the pragmatic turn）、向社會文化轉（the socio-cultural turn）、向外圍轉（the dispersive turn）、向經驗轉（the empirical turn）、向數位轉（the digital turn）、向口語轉（the discursive turn）以及向歷時轉（the diachronic turn）。這些轉向與二十世紀末以後人們開始用電腦的運算能力處理大量資料有關，在整個人文社會領域都經歷了向運算轉（computational turn）的大環境下，各個學科都開始有了研究方向上的巨大變動（Berry, 2011），就連傳播的閱聽

⁴ 語構學（syntactics）、語意學（semantics）、語用學（pragmatics）為國內常用的翻譯，亦可以依序翻譯成符號關係學、符號意義學與符號實用學，如劉慧雯、柯籙晏（2016 年 6 月 19 日）即參考趙毅衡（2011）的建議以「符形學」、「符意義學」與「符用學」翻譯，不過本文並不會特別探究符號與意義之間的關係，因此僅以最常用的翻譯為主。

人研究也開始從「語言學轉向」到「實踐轉向」（王宜燕，2012），於是更貼近生活、更注重實踐的各種研究方法論逐漸開始被研究者使用。

為了處理更大量的語言資料，結合語料庫語言學（corpus linguistics）和語用學的「語料庫語用學」（corpus pragmatics）的分析方式也就因應而生。過去在分析語言的用法時，研究者通常會順著文本的敘事順序以線性閱讀資料，Aijmer & Rühlemann（2015）將其稱之為水平閱讀方法論（horizontal-reading methodology），這樣的研究取徑更關注詞彙在文本的脈絡中展現出的意義，能做出的詮釋與分析也比較偏向質性研究；然而當欲分析的文本龐大到需要使用語料庫觀察時，就沒有辦法再靠人工逐一線性閱讀，而是必須透過工具將要觀察的字提取出來觀察，即為垂直閱讀方法論（vertical-reading methodology），然而這種方式的短處就是有時會因為把詞彙抽離脈絡而失真。而結合質性與量性分析的語料庫語用學（corpus pragmatics）就可以整合水平閱讀方法論與垂直閱讀方法論（integrated-reading methodology），在實際操作研究的過程中，可以利用語料庫分析常見的詞頻（frequency）、索引（concordance）、共現（collocation）等不同的功能來兼顧質性與量性的觀察（Aijmer, 2018）。

原先在語料庫語言學中，通常習慣從語言的格式、結構開始探索，進而分析格式與結構所反映的實用意涵，這樣從形式到功能（Form-to-Function）的研究取徑通常適用在比較小的語料庫，也跟前文的水平閱讀方法論更相似。然而隨著分析工具的進步，也可以實現先找出想研究的功能詞彙進而分析其展現形式（Function-to-Form）的研究取徑（Aijmer & Rühlemann, 2015），這種方式則比較接近前文的垂直閱讀方法論。這兩種研究取徑也各有其潛力與侷限，語言學家 Anne O’Keeffe 認為在語料庫研究的範疇中，研究者時常以犧牲脈絡的深度作為代價換

取形式上的廣度（*Breadth of form at the expense of contextual depth*）（O’Keeffe, 2018），指出在語言學的領域中質性與量性的揉合仍有一定的難度。因此她以語料庫為基礎，整理了前人「形式到功能」的研究取徑，並提出「功能到形式」的研究取徑該如何做到近乎質量並重，分別是（1）透過詞語標注便於研究者逐個搜尋需要的項目、（2）透過定義篩選原則縮小研究語料、（3）以現有研究做為種子研究滾雪球式發展以及（4）使用不同特性的語料庫。第一個方式指的是透過將文本中的所有字詞都依照研究需求進行標注（例如標注詞性、功能等），就可以幫助研究者在提取資訊時更加精確，進而從標注的功能順利追溯到語言的形式（例如人們在道歉時會使用什麼形式的語句），第二個方式則是直接透過原則性的剔除不想觀察的文本，來減小研究語料的量（例如只觀察使用頻率高的詞語用法），第三種方式則是直接將現有的研究進行比較，最後則是指直接挑選不同功能的語料庫進行分析（例如選擇日常用語語料庫或者選擇商務用語語料庫）。然而語言學領域中的語料庫都是長期累積的資料，雖然有細緻的標注或明確的功能分治，但對於傳播研究來說可能不夠新，或者方向不合用，因此就需要透過自建語料庫來進行語料庫分析。

上述的語料庫基礎、語料庫驅動、形式到功能、功能到形式這些不同的研究取徑並不是互斥的，有時因為研究的需求也會互相結合。Johansen（2019）便將這些研究取徑整理到一個矩陣當中來說明其關係（詳見表 2）：

表 1：形式與功能、語料庫基礎與驅動的交互關係

	形式到功能 (Form-to-Function)	功能到形式 (Function-to-Form)
語料庫基礎 (Corpus-based)	預定義 (pre-defined) 觀察的形式，再探索其功能	預定義 (pre-defined) 觀察的功能，再探索形式怎麼實現這個功能
語料庫驅動 (Corpus-driven)	透過語料庫呈現的數據 (例如詞頻) 來確立形式，之後才觀察功能	透過語料庫呈現的數據來確立功能，之後才觀察形式怎麼呈現

資料來源：譯自 Johansen (2019) 的研究 Exploring the use of probes in a corpus pragmatic study of hedging strategies. *Nordic Journal of English Studies*, 18(1), 125.

Johansen (2019) 認為過去的研究固然有過去 O’Keeffe 提到深度與廣度的兩難，但如果結合語料庫驅動中功能與形式的兩個方向 (作者自己定義為形式₁到功能到形式₂，原文為 form₁-to-function-to-form₂)，就能夠發展出語言的「探測器」(probes)，本文將統一稱之為探測語，亦即能夠透過搜尋包含該詞彙的文字段落便能夠找出特定功能語料的工具。

通常在語料庫研究中提取特定詞彙，指的是如同 Aijmer & Rühlemann 所說的垂直閱讀，通常是將欲觀察的標的物置中以後觀察上下文的脈絡，藉此了解特定詞彙可能的用法，也就是透過提取語料庫中的某些詞彙，就能夠判斷文章的意義 (即表 2 中的形式到功能)，應用在語言教學或研究上時，就能幫助研究者感知到當文章出現某個詞彙時，其語意就可能為何。Johansen 則是從語料庫語用學的角度再往前推進一步，做到「形式₁到功能到形式₂」，以下將詳細說明其作法。首先利用過去的文獻得知在挪威語的語境當中，每當人們在論述的過程中想要使用規避策略 (hedging strategies) 含糊其辭時，在英文的語境中常常

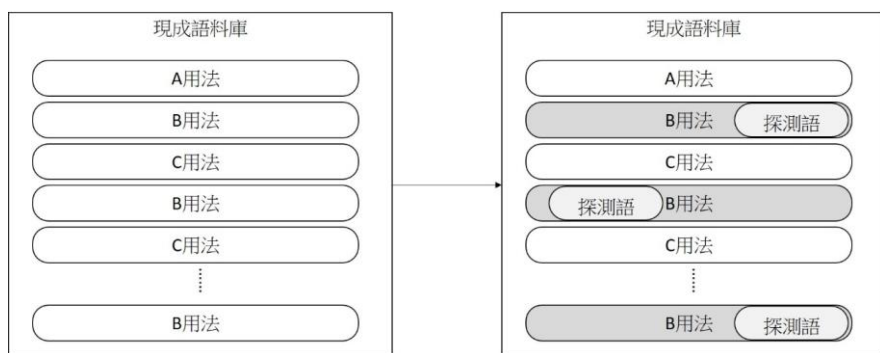
使用「but」，同樣在挪威語中的「men」也有相似的功能，⁵ 因此利用文獻就完成了預定義形式（pre-define form），也就是「形式₁」，接著透過提取語料庫中包含這兩類字的語料進行規避策略的功能分析，也就是第二步往功能的分析，最後再分析這些撈取出來的語料是否有共通的樣貌，也就是「形式₂」，最後結論得出了在挪威語的語境中，確實可以利用挪威語的「men」或英文的「but」這兩類詞作為規避策略的探測語。

若以將此邏輯應用至本研究中，本文試圖將鄉民眼中所見的意見氣候線索具象化，除了介面提供的線索以外，另一個 AUR 便是眾多其他鄉民在留言處累積堆疊的文字留言。因此如果想要知道當哪一些詞彙時常映入瀏覽者的眼簾時就會釋出「這是正在形成的意見氣候」這樣的線索時，就需要利用 Johansen 所提出「形式₁到功能到形式₂」的做法來找到專屬於意見氣候的探測語。Johansen 的研究對於本文的推進有相當大的助益，如果說 Johansen 研究中的核心關懷是「規避策略論述的探測語」，本文想提出的便是「意見氣候形成的探測語」，不過最大的差別在於 Johansen 在研究中作為探測語的詞彙是透過前人的文獻直接定錨於「men」與「但是」這兩類詞（如下圖 1），本文則是需要透過語料庫驅動的研究取徑（利用詞頻、索引等分析方式）找出探測語為何（如下圖 2）並檢視其預測功能，以及探討這樣的方式在方法論上是否有利用厚數據的特性達成質量並重的觀測方法設計，同時也建構觀測意見氣候的分析架構，讓探測語不需要局限於事先設定的詞彙，而是可以讓研究

⁵ 原文舉例英文利用「but」執行模糊策略的句子為 John is a Republican but he's honest，而用挪威語「men」執行模糊策略的句子 Det var ikke det jeg mente da, men，翻譯成英文是 It wasn't what I meant then, but，以中文來理解就像是「我不是那個意思但……」這樣還沒有說完但也不打算詳細解釋的含糊用法。

者持續地更新觀測。

圖 1：Johansen 的探測語用法



資料來源：研究者繪製

圖 2：本文的探測語用法



資料來源：研究者繪製

肆、研究發現

一、新媒體環境中準統計官能的樣貌

在文獻回顧中提到，閱聽人能夠從社群媒體上偵測到的社交線索包

含原著內容（proprietor content，以下簡稱 PC）、集體用戶表現（aggregate user representations，以下簡稱 AUR）以及使用者生產內容（user-generated content，以下簡稱 UGC）之上（Walther & Jang, 2012）。而 PTT 的 PC 與 UGC 是重疊的，文獻中的 PC 指的是例如新聞網站上的新聞報導、旅館網站的房型介紹等（Walther & Jang, 2012），便會與網友留下的評論 UGC 有區隔，PTT 為論壇網站，因此其 PC 為每一則鄉民的發文，雖然這些原生於 PTT 的文章並非由網站生產應該也要視為 UGC，但為了與下方大量累積的回覆作出區隔，故同樣視為 PC。而對鄉民來說，介面所呈現出的 AUR 是最主要影響注意力配置的要素。首先在開啟畫面時游標就會預設在「分組討論區」，因為所有 PC 與 UGC 的內容都會在分組討論區中。在分組討論區中如果有興趣進入特定主題（例如學校、動漫等）有初步的分類可以選擇，但最下方以雙箭頭強調的是「即時熱門看板」，對於想知道目前什麼事情最多人關注的鄉民，就會被引導去最熱鬧的地方。此外，在還沒有進入任何看板之前，就可以看到 AUR 的呈現，在每個板的中文敘述右方有一個標示「人氣」的欄位，呈現出前文提到的不同程度的人氣標示（如以紅字標示的「爆」等）。

此外不只是電腦版的介面有類似功能，PTT 相關的手機應用程式也都有類似的設計，例如最多人下載過的 MoPTT 以及次多人下載過的 Pitt 也都有直接展示出熱門看板與熱門文章的功能，因此閱聽人不論使用什麼載具打開 PTT，都可以在還沒有點開任何文章之前，就全觀地掌握目前平臺上「哪裡人多」。

接著進入本文撈取資料的八卦板後，AUR 的呈現就更直觀了，不論是每篇文章前方的「爆」或是搜尋「同標題文章」時會出現的一系列討論串，都可以非常直接地展現文章的熱門程度，推文數高展現出該篇

文章下方的討論熱度，相同標題陳列則顯示出議題的討論度越多相同甚至相似的標題出現，代表同主題的文章已經形成討論串。

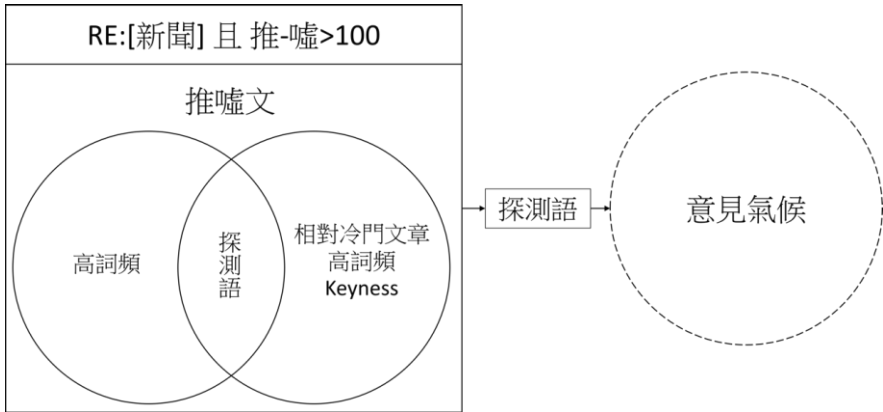
此外，轉載新聞的行為本身便隱含了閱聽人同時看向公領域與看向社群的視線，閱聽人接收了具有公共性的新聞資訊以後，帶著資訊轉身面對其他應該要看到這些資訊的他者，這樣的思維與準統計官能中閱聽人以直覺揣測他人對議題看法的思維非常接近，兩者都是在考慮「別人對於特定事件會怎麼想」，只是過去沉默螺旋的研究中是由研究者端著新聞資訊問受訪者認為主流意見氣候為何，本文觀測標題包含「Re: [新聞]」的文章以及下方的推噓文，則反映出閱聽人自己端著新聞面向他人所獲得的回應。

故依照介面與轉載新聞行為的邏輯，本文才會選擇推爆、Re:[新聞]的交集，因為依照鄉民進板的視線，包含這兩個特徵的文章代表了高人氣與高討論度的資訊。綜合上述的進板歷程，PTT 的介面讓閱聽人彷彿鳥瞰平臺上意見氣候的衛星雲圖，不像是使用 Facebook、IG 等以個人為中心的社群媒體只能看見以好友圈與追蹤內容構成的過濾泡泡，在 PTT 的使用歷程中，介面便不斷引導閱聽人往熱鬧的地方靠攏，而 AUR 的呈現則提供了閱聽人啟動準統計官能的重要社交線索。

二、依照準統計官能的資料蒐集邏輯：探測語

除了介面與 AUR，所有 UGC 的內容同樣也是觸發閱聽人準統計官能的要素，本文以一系列的步驟找出當月份的探測語，而彙整過後的成果也能夠觀察出探測語的細微變化。因此尋找探測語的步驟彙整如下圖 3：

圖 3：探測語尋找步驟與後續探測與分析方式



資料來源：研究者繪製

在這一系列找出探測語的步驟中，本研究實踐了表 2 中右下角功能到形式（Function-to-Form）且語料庫驅動（Corpus-driven）的文本探索，不同於 Johansen（2019）以特定的探測語抓取出語料庫中特定用法的語料，本文利用 AUR 的介面邏輯來確立探測語的實際功能，之後才透過自建語料庫並利用語料分析工具計算後呈現的數據觀察探測語形式怎麼呈現。

本文嘗試透過語料庫驅動的研究取徑找出探測語為何，更進一步地說，本文想找到的是能夠過濾出熱門議題的探測語為何，因此透過介面邏輯找出鄉民眼中熱門的貼文與下方的推文。

首先將下方大量的推文計算出詞頻（frequency），各月份詞頻前十高的排序如下表 3，高詞頻的詞彙代表的是依照介面邏輯，鄉民會看到的熱門文章下，推噓文當中最常出現的字眼。可以看到各月份都有非常類似的用詞，例如「推」、「笑死」、「塔綠班」等。不過這僅能呈現出熱門議題之下人們最常回覆這些詞彙，並不能排除鄉民在冷門議題之

下的回覆習性，因此本文也蒐集了冷門文章之下的回覆，並比較其詞與顯著程度，藉此檢證哪些詞彙僅出現在熱門文章之下。

表 3：2022 年一月到六月推文詞頻

一月		二月		三月		四月		五月		六月	
word	freq.	word	freq.	word	freq.	word	freq.	word	freq.	word	freq.
推	4,024	推	3,404	推	2,404	推	2,807	推	4,864	推	3,932
加油	1,476	垃圾	865	笑死	907	塔綠班	1,237	塔綠班	2,601	苗	2,121
笑死	1,394	塔綠班	857	臺灣	891	死	1,196	笑死	2,206	不是	1,744
塔綠班	1,351	笑死	845	塔綠班	802	笑死	938	民進黨	2,010	笑死	1,553
不是	1,191	臺灣	838	不是	761	高利貸	875	綠	1,992	塔綠班	1,342
臺灣	1,018	吃	711	民進黨	727	垃圾	859	不是	1,928	出來	1,326
政府	908	不是	676	黨	684	臺灣	846	垃圾	1,891	想	1,242
垃圾	853	一堆	603	垃圾	653	黨	837	黨	1,884	現在	1,164
才	738	政府	555	電	595	政府	824	臺灣	1,624	臺灣	1,090
想	717	想	554	現在	580	不是	777	現在	1,588	黨	1,084

資料來源：研究者彙整

同樣依照介面呈現的邏輯，本文蒐集冷門文章的定義為「回覆不超過 10 則」的文章。根據 PTT 的介面，當推文減噓文的數量不超過 10 的時候，文章前方會顯示綠色的 1 到 9 不等的數字，超過 10 的話則會變成黃色，超過 100 則會顯示爆。為了避免有大量回覆但是相減仍然為 10 以內的文章出現，本文選擇回覆不超過 10 則的文章，如此在文章前方只會顯示綠色的 1 到 9 或是只顯示灰色的「-」，詳細的數據則可以參考表 4。

表 4：冷門文章資料蒐集則數綜覽

2022 年月份	主文	推文
一月	13,460	71,242
二月	13,118	69,471
三月	12,505	59,464
四月	11,121	53,462
五月	11,125	54,073
六月	9,727	46,979
合計	71,056	354,691

資料來源：研究者彙整

透過將熱門文章的回覆與冷門文章的回覆比較，本文進行關鍵字詞計算（keyness calculation），藉此檢證哪些詞彙是相較於冷門文章，在熱門文章才更容易出現的常用回覆。這種計算的方式類似於 TF-IDF（Term Frequency-Inverse Document Frequency），是為了找出大量文本中重要詞彙的方式。

TF-IDF 是一種基於詞頻和文檔頻率的加權計算方法。它結合了詞語在文檔中的頻率（TF）和在整個語料庫中的逆文檔頻率（IDF）。TF-IDF 的目的是強調在特定文檔中頻繁出現但在整個語料庫中相對較少出現的詞語，這些詞語被認為對於該文檔的主題或內容具有重要性。TF-IDF 最早由 Robertson & Jones（1976）提出，後續也經過幾次修正，多用於搜尋引擎的結果呈現。TF-IDF 的概念為了削弱 Shannon 提出資訊理論（Information Theory）中的「熵」（entropy），亦即雜訊，因此在計算的邏輯上將文字頻率大的單詞定義為越無用（Robertson, 2004），這樣的計算方式放在本文的研究中並不合理。若利用 TF-IDF 會將每篇熱門文章中都出現的回覆視為不重要，反之只出現在少數文章中的詞彙

才重要，便悖離了本文想找出熱門文章探測語的初衷，因此本文選擇使用關鍵字詞計算（keyness calculation）。

關鍵字詞計算（keyness calculation）旨在評估詞語在一個文檔或文本集合中相對於整個語料庫的數值重要性（Scott, 2010）。Keyness 測量的目的是判定詞語在文檔中的使用是否與預期相符，是否與整個語料庫中的使用情況有顯著差異。通常使用的 Keyness 測量方法包括卡方檢驗（Chi-square test）、對數似然比（log-likelihood）等，這些測試可以計算出詞語的關鍵性指標，該指標表示詞語在文本中的使用是否統計上顯著。

本文利用庫博語料分析工具（Corpro）中的詞語顯著性計算功能，直接排序出比較熱門文章推文與冷門文章推文的 keyness 數值，將其中值為正向的詞彙整理成表 5。

表 5：與冷門文章推文比較之正向顯著常用詞彙

	一月	二月	三月	四月	五月	六月					
word	keyness	word	keyness	word	keyness	Word	keyness	word	keyness	word	keyness
推	1977.9	推	2057.95	推	1111.70	推	1125.26	推	1018.14	推	843.84
加油	1012.2	高調	400.51	綠能	317.92	高利貸	631.85	塔綠班	485.29	苗	781.51
塔綠班	518.3	垃圾	348.93	笑死	293.82	塔綠班	467.87	垃圾	380.62	笑死	254.28
笑死	499.1	塔綠班	280.73	塔綠班	282.87	全家	452.72	笑死	353.46	塔綠班	249.06
消防人	335.6	笑死	246.88	垃圾	234.30	垃圾	258.48	綠	352.58	綠能	225.90
重視	309.5	死忠的	212.05	民進黨	215.98	笑死	249.13	綠能	278.77	議員	221.25
需求	290.2	專業	161.67	黨	178.88	粗暴	229.76	柯	276.04	網軍	201.07
台北通	258.6	中天	134.72	記者	173.32	死	204.03	高調	250.84	雙標	179.93
政府	209.5	政府	128.43	綠共	156.30	柯	186.61	投	229.23	呵呵	164.37
柯	190.6	粗暴	125.62	死忠的	145.74	綠能	181.89	辛苦	223.80	噁心	160.57

資料來源：研究者彙整

表 5 中所呈現的，都是相較於冷門文章推文，更常出現在熱門文章下的詞彙。因此如果取表 3 與表 5 的交集，篩選出既確保是熱門文章才會出現，又是鄉民很常使用的詞彙，也就是本文想尋找的探測語，整理如下表 6。

表 6：各月探測語

2022 年月份	探測語
一月	推、加油、塔綠班、笑死、政府
二月	推、垃圾、塔綠班、笑死、政府
三月	推、笑死、塔綠班、垃圾、民進黨、黨
四月	推、高利貸、塔綠、垃圾、笑死、死
五月	推、塔綠班、垃圾、笑死、綠
六月	推、苗、笑死、塔綠班

資料來源：研究者彙整

三、以準統計官能設計之輿論觀測法的預測性

由於意見氣候變化是一個速度很快且持續的過程，在觀測上需要能夠突破單次觀測的框架，本文想得知閱聽人在當代的媒介環境下利用準統計官能是否可以大致勾勒出正在形成的意見氣候，便以每個月的探測語作為關鍵字，搜尋下個月推文中含有探測語的文章，並計算這些文章的爆文數量。選擇以爆文數量作為觀測標準是因為這是最直觀且易於篩選的 AUR，並且為了確認其預測效能，也列出該搜尋時間段中的爆文總數作為比較（見表 7），可以看到探測語預測出下一個月首週大部分的爆文。

表 7：探測語預測結果

預測驗證	爆文則數/爆文總數	預測比例
一月探測語預測二月首週	236/240	98.3%
二月探測語預測三月首週	421/426	98.8%
三月探測語預測四月首週	295/298	99.0%
四月探測語預測五月首週	402/403	99.8%
五月探測語預測六月首週	313/317	98.7%

資料來源：研究者彙整

表格中「爆文則數／爆文總數」前者指的是「探測語抓出的爆文／當月爆文總數」，透過與撈取本文資料相同的工具，以一月的探測語為例，撈取二月第一週推文中包含「推、加油、塔綠班、笑死、政府」其中任何一個以上一月探測語的文章，並只篩選出推爆的文章觀察。為了檢視探測語抓出的爆文數量與爆文總數的差距，也利用同樣的工具計算第一週有多少則爆文，從比例上可以觀察到探測語幾乎能夠抓出所有熱門文章，以五月預測六月首週的爆文為例，其中四篇不是由探測語抓出的文章如下表 8：

表 8：不包含探測語的爆文列表

文章代碼	文章標題	count
M.1654068318.A.0A1	[問卦] 喜迎家裡第三隻小貓 幫取名	257
M.1654138643.A.5F9	[新聞] 陳道明驚傳過世享年 66 歲! 張信哲、伍思凱	190
M.1654224423.A.BD2	[問卦] 認真問 還沒染疫的人+1	244
M.1654474108.A.F46	[問卦] 你們怎稱呼 7-11 ?	173

資料來源：研究者彙整

第一篇幫小貓取名的文章下方留言都是網友在隨意提供寵物名字或者政治人物名字；第二篇是大家表示以為過世的是飾演康熙皇帝的演員王道明，推文不斷重複皇上駕崩，結果過世的是資深音樂人王道明；第三篇是許多人推「+1」；第四篇則是鄉民們任意地回答店家的名字，包括 Costco 以及 IKEA。這些文章的特徵都是大家沒有在討論事情，只是反射地回應一些很短的詞彙，並沒有辦法展現出任何意見氣候特徵（諸如文獻與數據結果提到的跨平臺溢散、新生成資訊等）。因此，透過探測語能夠排除掉這類玩鬧且沒有生成意見氣候的爆文，只抓出形成中的意見氣候。

預測結果中逐月的內容又在討論什麼，本文則利用詞頻與共現組合展示出當週的意見氣候樣貌。透過詞頻繪製出的文字雲可以直覺地看到聲量大小，利用各個詞彙之間共現的 T-score 繪製成的網絡圖則可以看到議題之間如何被談論以及哪些議題會伴隨著出現。

由於探索未知事件時並不知道事件週期的長短，故以週為單位繪製文字雲。二月的第一週從文字雲觀測到的意見氣候大多和疫情有關，例如文字雲中間字體比較大的「例」以及周圍比較小的「確診」、「本土」都是轉載新聞報導當日確診數會出現的字眼，完整的句子通常會像是「高雄今（4）日新增 12 例本土確診個案」，也因此會伴隨著出現許多地名。此外從詞頻還可以看到人名「黃郁婷」，是因為當時正在舉辦北京冬奧，臺灣代表隊的滑冰選手黃郁婷在 Instagram 上傳了一張穿著中國隊服的照片，被新聞報導後在 PTT 上也受到了許多討論，因此也能看到相關的詞彙包括「選手」、「北京冬奧」。三月的第一週則抓出了兩個最大的事件，包括全台大停電以及俄烏戰爭（圖 5），四月的首週則又回到討論疫情相關的主題內容（圖 6）。

此外，光是從探測語在逐月之間的微小變化，也能夠反映意見氣候轉變的過程，例如當月如果有討論度特別高的議題，就會出現在探測語當中，這意味著閱聽人除了透過介面與數字展現的 AUR 之外，在瀏覽 UGC 的過程會更常看到這些詞彙，對於閱聽人來說就可以意會「大家都在談論這個」，就像四月的「高利貸」與六月的「苗」，都是因為當月有相關的熱門議題。

伍、討論與結語

本文的研究目的在補足過去沉默螺旋理論中對於準統計官能原理的描述，並利用介面原理建立一種能夠跨時間、跨議題且更貼近鄉民視線的意見氣候分析框架，主要的研究貢獻分別為勾勒當代媒介環境中閱聽人的準統計官能運作原理以及提出意見氣候分析的新分析框架，詳述如下：

一、勾勒當代媒介環境中閱聽人的準統計官能運作原理

回到一開始提及的沉默螺旋理論，Noelle-Neumann 在二次戰爭的時空背景下擔憂強勢意見會掩蓋沉默多數的聲音而提出了沉默螺旋理論。過去理論對於閱聽人的行為描寫，從利用準統計官能偵測意見氣候開始，到最後決定是否公開表達自己逆風的意見，一直都有部分無法解釋的細節，本文透過重新審視當代媒介景觀補足了理論未能解釋之處。從上述的數據結果可以推論，過去未能說明原理的準統計官能確實存在，並且在當代媒介環境下能夠理解為：「閱聽人透過接收 PC、AUR 與 UGC 所釋出的社交線索拼湊整體意見氣候的連續行為」。至於後續是

否會落入沉默的螺旋，則因為時空背景與媒介景觀的不同，而有了沉默與否以外更多元的選擇。

本文專注於釐清在新媒體環境中準統計官能的原理如何運作，首先從文獻得知閱聽人確實有透過資訊介面拼湊意見氣候的行為，並在「研究發現」裡檢視撈取到的資料，盡可能的印證文獻中提到的內容。

二、提出意見氣候分析的新分析框架

了解當代的準統計官能原理則是本文的另一學術貢獻，即揭示多元資訊環境下意見氣候分析的新途徑。利用語料庫語用學的觀點出發，為了找到「形式₁到功能到形式₂」中，發揮提取意見氣候功能讓我們看到熱門意見樣貌的「形式₁」，也就是本文要找的探測語，本文透過一連串的步骤萃取出每一月份的探測語，並且回頭驗證也發現探測語確實能夠發揮一定的預測功能。

此外，透過介面邏輯找出熱門文章下方推文的共有用語，想藉此歸納出具有預測能力的探測語，本文發現每月份的探測語在一段時間內大部分會相同，而且對於下一個月出的熱門文章預測也有一定的效果。在文獻中提到，本文希望透過語料庫驅動的研究取徑（利用詞頻、索引等分析方式）找出探測語為何，以及探討這樣的方式在方法論上是否有利用厚數據的特性達成質量並重的觀測方法設計。結果顯示透過這樣的方式比起直接撈取爆文，更能夠找到鄉民真正透過留言累積構築而成的意見氣候，而非其他種類的爆文，例如沒有意見氣候累積的爆文（例如前述的小貓取名文）或問卷等。本文提出新分析框架的價值在於：

（一）跨出預設的研究範圍

本文以資料先行的研究取徑完成一次探索式的研究，一方面是為了更忠實地呈現意見氣候生成的過程，另一方面則是想建立一個跨越主題的語料庫分析研究框架。首先關於忠實地呈現意見氣候生成的過程，本文因為從資料蒐集就改變過去先限定主題或先下關鍵字的蒐集方式，並且考量到介面的作用力藉此重現「鄉民眼中所見」，這樣的研究取徑讓研究者的真實貼近閱聽人眼中的真實。進入 PTT 看熱鬧的鄉民看不到後台數據，也看不到個別的帳號之間會形成什麼網絡，更不會只看見選定好主題的文章，研究者若完全致力於追求大數據的數量以及肉眼無法觀測到的數據關係，則會形成閱聽人永遠無法觸及的真實。本文試圖見鄉民之所見、聞鄉民之所聞，讓研究者與閱聽人的視線重合，觀測到更貼近閱聽人眼中的意見氣候。

其中，探測語是沿著鄉民的視線逐步蒐集、整理與剔除後所得的結果，而探測語在使用者實踐上的意義，代表當鄉民看到這些推文的字眼，通常會認知到正在瀏覽的是群眾熱議中的文章。而後續意見氣候的預測與蒐集，則可以透過蒐集回覆包含探測語的文章與推文進行觀察，亦可更進一步利用語料分析工具計算與探測語共現的詞彙來細究群眾熱議的議題內容。以這樣的方式尋找意見氣候便可以突破輿論分析時僅能限制在研究者想搜索的議題，且分析方式難以在不同議題之間通用的困境。

（二）可永續使用的分析取徑：機動式探測語

Noelle-Neumann 在提出沉默螺旋之前就已經指出過，研究媒介效果需要呈現出累積的過程，所以應該採用小樣本連續調查（panel）搭配複式方法（muti-method approach）與相關的外在資料來實證媒介效果（翁

秀琪，1990）。然而不論是沉默螺旋原始的資料蒐集方法，或是當在透過資料探勘以議題切入的資料蒐集與分析方式，都是一次性的操作，觀測的內容隨著選情結束嘎然而止，又或者是探勘的資料分析結果不能通用到其他主題之上，同樣隨著議題結束而不再能夠記錄意見氣候變化的過程，因此本文提出的機動式探測語將初步地突破上述限制，讓研究者能夠對意見氣候以合理可循的架構進行任意時間長度的觀測。

此外，從語料庫語用學的觀點，本文也將探測語的分析方式再次改良，奠基於 Johansen（2019）的研究上，本文將固定的探測語設計成能夠隨著觀察時間推移而持續更新、擴充的詞組，且初步地測試出預測功能，讓原本只能固著在既定語料庫中的「形式₁到功能到形式₂」（form₁-to-function-to-form₂）能夠突破界線，不再只能以預先設定好的關鍵字觀察，而是從持續累積的意見氣候場域中將持續新生成的資訊自由地納入自建語料庫中。

三、研究限制

（一）未能涵蓋其他社群介面的作用力

本文選擇 PTT 作為觀察對象，因為能夠排除研究者無法掌握的平臺演算法，觀察介面如何以最直觀的方式引導閱聽人的視線，然而不同的媒介的介面與功能會以不同的方式引導閱聽人趨向熱鬧的地方，例如影音平臺 YouTube 首頁就在左方的工具列提供「發燒影片」的按鈕，讓閱聽人可以直接瀏覽近期觀看次數最多的影片，或者 Instagram 會直接利用應用程式的推播功能通知使用者「快來看看今天在 Taiwan 觀看次數最多的連續短片」，引導閱聽人直接去最多人聚集的地方，因此介面的作用力在不同的媒介中會發揮不同的引導力量。

不過以社群為核心的媒體就不一定會由介面鼓勵閱聽人聚集，以 Facebook 為例，在閱聽人眼中打開 Facebook 所瀏覽的是以好友動態及追蹤的粉絲專頁為大宗構成的資訊內容，其中也穿插著依照使用者偏好而推播的「為你推薦」內容（包含粉絲專頁、公開社團等）；同樣 Instagram 除了上述的「快來看看今天在 Taiwan 觀看次數最多的連續短片」，也有「建議貼文」的內容（演算法認為你喜歡並建議你追蹤的公開 IG 帳號），這些透過演算法直接將資訊推送到閱聽人眼前的個人化機制對本研究而言，如同遮蔽閱聽人感知意見氣候的準統計官能，在蒐集社交線索的過程中，可以近用的 AUR 因為個人化機制而變得破碎，於是每一位閱聽人眼中的「大家」都不是同一群人，若沒有特別有意識地從不同介面機制的媒體蒐集多元資訊來源，閱聽人很容易就會被困在過濾泡泡（filter bubble）當中。即便是本文選擇較沒有演算法影響的 PTT，在 2023 年也出現了「八又翻」一詞，為「八卦板又翻車」的縮寫，指涉的是八卦板鄉民時常以為事情或選情會如何發展最後卻不如預期而「翻車」，過去鄉民也時常自嘲八卦板為社會的「反指標」，因此認識閱聽人如何在不同媒體上檢索意見氣候也對於資訊素養的培養有重要的幫助。

（二）無法排除極端個案的影響

本文透過介面邏輯蒐集資料，是為了讓研究者的視線與鄉民的視線更加接近，而非利用後台資訊觀測出閱聽人永遠不會看見的真實，然而此作法只能盡可能地還原數量且在八卦板圍觀意見氣候的鄉民使用 PTT 時的視線途徑，並非指涉所有的鄉民在進板以後都同樣會追隨著 AUR 到最熱鬧的八卦板並且點選被推爆的文章，個別的鄉民也可能設定屬於他自己的「最愛看板」直接將視線投注於自己感興趣的社群中，

本研究的觀測方式確實無法用來指稱每一位鄉民的共通行為。

此外，有時候個別鄉民的行為可能會對整體數據結果產生影響。例如有些文章作者會在文末指定回覆特定文字，依照推文順序挑選幾個回覆發放平臺的虛擬貨幣「P 幣」，藉此讓自己的文章更容易被看見，類似這樣刻意操作影響 AUR 的行為偶爾會使得數據被改變，但若拉長觀測時間或是設定過濾條件，即可減小這種情況造成的影響。

本文觀測的平臺批踢踢實業坊在 2018 年 9 月 14 日深夜發布公告，為了解決浮濫註冊帳號的問題停止新帳號註冊，2019 年 5 月起至 11 月之間陸續刪除上千個異常帳號，2020 年 3 月 4 日開放擁有臺大信箱的民眾註冊帳號，直到 2021 年 7 月 26 日才完全開放給所有人註冊使用，歷經這樣的變革使得 PTT 上的使用者成分受到影響，但本研究未能進一步比較變革前後的意見氣候是否有顯著差異。

更甚者，本研究無法精確分辨出熱門文章下方的回覆有多大比例是由公關公司等機構的操作所造成的。例如本文蒐集各個月份推文最多的文章觀測「起風的瞬間」只能觀測到意見氣候累積的速度，卻很難得知究竟起風的原因是突發事件引起網友們的圍觀，還是有對價關係的公關公司或網軍所為；又例如議員苗博雅引發的「推文朗讀事件」在短短三天之內風向就從討論事件轉變為人身攻擊，本文沒有辦法得知帶起人身攻擊風向的網友身分為何，因此也沒辦法指認風向是否為網軍所影響。然而，從鄉民的觀點來看，無論留下回覆的帳號持有者是否與特定勢力存在對價關係，研究者蒐集到的資料與鄉民看見的實情並無二致，在本研究的範圍中確實體現了「鄉民眼中的意見氣候」，然而未來若要觀測更貼近線下實情的意見氣候，甚至更進一步地探究這些鄉民眼中的意見氣候在現實世界中究竟反映了 Noelle-Neumann 在沉默螺旋理論中提到的多數意見還是其實只反映出強勢意見，則需要結合其他資料蒐集與分析

的方式才能做出更進階的觀察。

四、未來研究發展建議

(一) 展望其他平臺的比較研究

此外，透過探測語構聯出 PTT 上的熱門議題除了能讓研究者一窺平臺上鄉民正在形成什麼意見氣候，未來若要繼續推展至其他平臺，亦可以相同的方式比較不同平臺之間意見氣候的樣貌差距。例如本文測試以一月探測語預測二月第一週的話題，從文字雲中可以看到比例不高但仍然可見的「黃郁婷」，若以 Google Trend 觀測同樣時間段人氣竄升的搜尋內容，「黃郁婷」則排名第一，從此可以看到幾個要點，例如探測語預測出的話題多少能夠與平臺以外的世界相符合，以及鄉民的意見氣候與平臺之外的意見氣候差異為何，更進一步還能討論彼此介面作用力的不同。例如 Google 搜尋引擎鼓勵人們搜尋有興趣的話題，PTT 則是引導大家聚集到熱門的話題之下，故獲得的資料也不同；以遊戲為主的平臺巴哈姆特（gamer.com）的介面以不同的遊戲作為分野，但進入討論板以後同樣也有與 PTT 類似的功能，讓瀏覽的網友知道目前哪一些文章比較熱門。Dcard 可以讓使用者選擇自己喜歡的看板追蹤，但同樣可以在所有看板的列表介面找到「即時熱門看板」，讓隨意瀏覽的閱聽人知道目前人們都聚集在哪裡。

有鑑於若干熱門的平臺都有引導閱聽人往熱門處聚集的功能，未來便能夠以本文的分析途徑比較不同的平臺，甚至可以比較國內不同的社群平臺甚至國際之間不同平臺的意見氣候樣貌，藉此一窺不同資訊生態中意見氣候的行程有何差異。不過未來需要注意的是，本文選擇了沒有演算法影響的 PTT 作為分析架構的雛型，PTT 以外的社群平臺不可避免

地會需要將演算法在介面中扮演的角色也納入討論。

此外，若選擇其他平臺進行觀察，依照不同的介面也需要調整資料篩選標準，例如本文的「RE: [新聞]」是針對 PTT 平臺的資料結構以及轉載新聞的行為作為蒐集資料的篩選條件，然而其他平臺並不會有「RE: [新聞]」這樣的資訊結構，就需要視情況調整。例如 Twitter、Facebook、Instagram、YouTube 都有的「hashtag」功能，也是適合觀測大量意見氣候的重要功能，未來的研究者可以參考與 hashtag 功能相關的文獻來決定適合的篩選標準。

（二）洞悉鄉民視角作為培養資訊素養的根基

為什麼需要了解閱聽人眼中的意見氣候，除了想找到跨越主題的研究設計框架，本文認為瞭解閱聽人的準統計官能機制在資訊識別的功能上也具有正面的意義，美國大學和研究圖書館協會於 2015 年提出的《高等教育資訊素養架構》（Framework for Information Literacy for Higher Education）中明訂了「資訊素養」的六個核心概念，包含權威是被建構且有脈絡的（Authority Is Constructed and Contextual）、資訊產製是一持續的過程（Information Creation as a Process）、資訊本身是有價值的（Information Has Value）、研究即調查（Research as Inquiry）、學術即對話（Scholarship as Conversation）以及檢索為有策略的探索（Searching as Strategic Exploration），架構中提到不論是研究或是閱聽人的檢索行為，都是持續的過程。然而在資訊爆炸的媒體環境下，持續有策略地檢索或是持續調查都有一定的難度，因此發展能夠不受限於特定時間或議題，且能夠以同樣分析框架持續蒐集資料的意見氣候觀測方式，不論對研究者或是媒體的使用者都能有幫助。對研究者來說，在觀測意見氣候時可以做到更長期且跨議題的觀測；對使用者來說，探測語

是精簡了大量瀏覽後才能獲得的社交線索，看到探測語即可意會到意見氣候正在形成，進一步地針對資訊判讀也會更有效率。

參考書目

- 王宜燕（2012）。〈閱聽人研究實踐轉向理論初探〉，《新聞學研究》，113，39-75。https://doi.org/10.30386/MCR.201210_(113).0002
- 朱景鵬、朱鎮明（2006）。〈政府效能對國際競爭力影響評析〉，《研考雙月刊》，30(6)，6-23。https://doi.org/10.6978/YKSYK.200612.0006
- 李柏彥、劉慧雯、紀明德（2021年6月25-27日）。〈英雄所見略同？社群媒體使用者共現行為之研究〉【論文發表】。「2021年中華傳播學會年會」，臺北市，臺灣。
- 林思平（2017）。〈網路八卦與真理政治：批踢踢八卦板之閱聽人研究〉，《新聞學研究》，133，135-188。https://doi.org/10.30386/MCR.201710_(133).0004
- 紀慧君（2018）。〈從語料分析探究有機食物之媒體再現：三十年之意義與轉變〉，《中華傳播學刊》，34，209-252。https://doi.org/10.3966/172635812018120034006
- 孫秀蕙（1993）。〈關於沈默螺旋的一場筆戰－淺談意識型態與理論建構的關係〉，《新聞學研究》，47，107-115。https://doi.org/10.30386/MCR.199306_(47).0007
- 孫秀蕙（1994）。〈民意理論中的「死硬派」之研究－以「興建核四廠」議題為例〉，《新聞學研究》，49，157-190。https://doi.org/10.30386/MCR.199407_(49).0008
- 徐士傑、邱兆民、洪郁雯、陳儷慧（2013）。〈影響個人提供線上社會支持因子之探討：以 PTT 癌症板為例〉，《中山管理評論》，21(3)，511-544。https://doi.org/10.6160/2013.09.02
- 翁秀琪（1990）。〈民意與大眾傳播研究的結合－諾爾紐曼和她的「沉默的螺旋」理論〉，《新聞學研究》，42，71-101。https://doi.org/10.30386/MCR.199001_(42).0005
- 翁秀琪（1997）。〈選民的意見形成－以民國八十二年臺北縣縣長選舉為例檢驗「沉默螺旋理論」〉，《新聞學研究》，55，160-182。https://doi.org/10.30386/MCR.199707_(55).0010
- 郭文平（2015）。〈字彙實踐及媒體再現：語料庫分析方法在總體經濟新聞文本分析運用研究〉，《新聞學研究》，125，95-142。https://doi.org/10.30386/

MCR.201510_(125).0003

- 傅文成、陶聖屏（2018）。〈以大數據觀點探索網路謠言的「網路模因」傳播模式〉，《中華傳播學刊》，33，99-135。https://doi.org/10.6195/cjcr.201806_33.0003
- 曾虛白（1969）。〈民意測量質量之辯〉，《新聞學研究》，4，1-5。https://doi.org/10.30386/MCR.196912_(4).0001
- 黃厚銘、林意仁（2013）。〈流動的群聚（mob-ility）：網路起關的社會心理基礎〉，《新聞學研究》，115，1-50。https://doi.org/10.30386/MCR.201304_(115).0001
- 黃厚銘（主編）、黃上銓、林意仁、蕭煒馨、李紹良（2016）。《婉君妳好嗎？給覺醒鄉民的 PTT 進化史》。群學。
- 楊意菁（2013）。〈網路意見的新聞再現與公眾想像：「網友說」新聞的內容與論述分析〉，《中華傳播學刊》，24，119-164。https://doi.org/10.6195/cjcr.2013.24.04
- 廖春閔（2018）。《應用集合視覺化之社群網路使用者交集活動探索與分析工具》。國立政治大學資訊科學系碩士論文。【國立政治大學博碩士論文全文影像系統】https://doi.org/10.6814/THE.NCCU.CS.019.2018.B02
- 趙玉娟、陶振超（2018）。〈政治網路口碑的情感分析：語意關連性之觀點〉，《傳播研究與實踐》，8(2)，75-107。https://doi.org/10.6123/JCRP.2018.07_8(2).0003
- 趙毅衡（2011）。《符號學原理與推演》。南京大學出版社。
- 劉慧雯（2017）。〈建構「倫理閱聽人」：試論社群媒體使用者的理論意涵〉，《新聞學研究》，131，87-125。https://doi.org/10.30386/MCR.201704_(131).0003
- 劉慧雯（2020）。〈從新聞到粉絲團：社群小編重構公共話語現象的初探研究〉，《傳播與社會學刊》，54，161-192。https://doi.org/10.30180/CS.202010_(54).0007
- 劉慧雯、柯籙晏（2016 年 6 月 19 日）。〈邁向厚數據：以「詮釋型資訊工具」進行意義分析的概念基礎〉【論文發表】。「2016 年中華傳播學會年會」，嘉義縣，臺灣。https://reurl.cc/xaQZgLhttps://reurl.cc/xaQZgL
- 鄭宇君（2017）。〈探討社交媒體事件之浮現邏輯：一個融合 STS 與傳播研究取徑之嘗試〉，《中華傳播學刊》，32，129-164。https://doi.org/10.6195/cjcr.2017.32.04
- 鄭宇君、陳百齡（2014）。〈探索 2012 台灣總統大選之社交媒體浮現社群：鉅量資料分析取徑〉，《新聞學研究》，120，121-165。https://doi.org/10.30386/MCR.201407_(120).0004
- 鄭芳婷（2016）。〈數位離散與酷兒想像：以拉板、彩虹山莊及 Queerology 為

- 例》，《考古人類學刊》，85，51-82。https://doi.org/10.6152/jaa.2016.12.0003
- Aijmer, K. (2018). 22. Corpus pragmatics: From form to function. In A. H. Jucker, K. P. Schneider, & W. Bublitz (Eds.), *Methods in pragmatics* (pp. 555-585). De Gruyter Mouton. https://doi.org/10.1515/9783110424928-022
- Aijmer, K., & Rühlemann, C. (Eds). (2015). *Corpus pragmatics: A handbook*. Cambridge University Press.
- Berry, D. M. (2011). The computational turn: Thinking about the digital humanities. *Culture machine*, 12. https://culturemachine.net/wp-content/uploads/2019/01/10-Computational-Turn-440-893-1-PB.pdf
- Buozis, M. (2019). Doxing or deliberative democracy? Evidence and digital affordances in the *Serial* subreddit. *Convergence*, 25(3), 357-373. https://doi.org/10.1177/1354856517721809
- Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The guardian*. https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election
- Chung, S. F., Ahrens, K., & Huang, C. R. (2005, December). Source domains as concept domains in metaphorical expressions. *International Journal of Computational Linguistics & Chinese Language Processing*, 10(4), 553-570. https://doi.org/10.30019/IJCLCLP.200512.0012
- Dai, Y., Van Der Heide, B., Mason, A. J., & Shin, S. Y. (2019). The wisdom of the crowd versus the wisdom in the crowd: Testing the effects of aggregate user representation, valence, and argument strength on attitude formation in online reviews. *International Journal of Communication*, 13, 3488-3511. https://doi.org/10.1016/j.jeconom.2020.07.016
- Davis, J. L., & Graham, T. (2021). Emotional consequences and attention rewards: The social effects of ratings on Reddit. *Information, Communication & Society*, 24(5), 649-666. https://doi.org/10.1080/1369118X.2021.1874476
- Dodge, S. (2021). A data science framework for movement. *Geographical Analysis*, 53(1), 92-112. https://doi.org/10.1111/gean.12212
- Donsbach, W., Salmon, C. T., & Tsifti, Y. (2014). *The spiral of silence: New perspectives on communication and public opinion*. Routledge. https://doi.org/10.4324/9780203125007-26
- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80(S1), 298-320. https://doi.org/10.1093/poq/nfw006
- Glynn, C. J., & McLeod, J. M. (1984). Public opinion du jour: An examination of the spiral of silence. *Public Opinion Quarterly*, 48(4), 731-740. https://doi.org/10.1086/268879
- Ingram, M. (2022, May 26). Facebook's new data-sharing plans raise old concerns.

- Columbia Journalism Review*. https://www.cjr.org/the_media_today/facebook-new-data-sharing-plans-raise-old-concerns.php
- Johansen, S. H. (2019). Exploring the use of probes in a corpus pragmatic study of hedging strategies. *Nordic Journal of English Studies*, 18(1), 121-148. <https://doi.org/10.35360/njes.493>
- Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology*, 31(2), 177-192. <https://doi.org/10.1002/ejsp.36>
- Knobloch-Westerwick, S., Sharma, N., Hansen, D. L., & Alter, S. (2005). Impact of popularity indications on readers' selective exposure to online news. *Journal of Broadcasting & Electronic Media*, 49(3), 296-313. https://doi.org/10.1207/s15506878jobem4903_3
- Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24), 8788-8790. <https://doi.org/10.1073/pnas.1320040111>
- Leavitt, A., & Robinson, J. J. (2017). The role of information visibility in network gatekeeping: Information aggregation on Reddit during crisis events. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, 1246-1261. Association for Computing Machinery. <https://doi.org/10.1145/2998181.2998299>
- Lee, B., Lancendorfer, K. M., & Lee, K. J. (2005). Agenda-setting and the Internet: The intermedia influence of Internet bulletin boards on newspaper coverage of the 2000 general election in South Korea. *Asian Journal of Communication*, 15(1), 57-71. <https://doi.org/10.1080/0129298042000329793>
- Lee, F. E. (2015). How party polarization affects governance. *Annual Review of Political Science*, 18, 261-282. <https://doi.org/10.1146/annurev-polisci-072012-113747>
- Lee, S., Atkinson, L., & Sung, Y. H. (2022). Online bandwagon effects: Quantitative versus qualitative cues in online comments sections. *New Media & Society*, 24(3), 580-599. <https://doi.org/10.1177/1461444820965187>
- Leong, A. D., & Ho, S. S. (2021). Perceiving online public opinion: The impact of Facebook opinion cues, opinion climate congruency, and source credibility on speaking out. *New Media & Society*, 23(9), 2495-2515. <https://doi.org/10.1177/1461444820931054>
- Morris, C. W. (1938). Foundations of the Theory of Signs. In O. Neurath (Ed.), *International encyclopedia of unified science* (pp. 1-59). Chicago University Press. <https://doi.org/10.2307/2267781>
- Neubaum, G., & Krämer, N. C. (2017). Monitoring the opinion of the crowd: Psychological mechanisms underlying public opinion perceptions on social media. *Media Psychology*, 20(3), 502-531. <https://doi.org/10.1080/15213269.2016.1211539>
- Noelle-Neumann, E. (1973). Return to the concept of powerful mass media. *Studies of broadcasting*, 9(1), 67-112.

- Noelle-Neumann, E. (1984). *The Spiral of silence*. University of Chicago Press. <https://doi.org/10.1177/016224398501000319>
- O’Keeffe, A. (2018). 23. Corpus-based function-to-form approaches. In A. H. Jucker, K. P. Schneider, & W. Bublitz (Eds.), *Methods in pragmatics* (pp. 587-618). De Gruyter Mouton. <https://doi.org/10.1515/9783110424928-023>
- Robertson, S. (2004). Understanding inverse document frequency: On theoretical arguments for IDF. *Journal of Documentation*, 60(5), 503-520. <https://doi.org/10.1108/00220410410560582>
- Robertson, S. E., & Jones, K. S. (1976). Relevance weighting of search terms. *Journal of the American Society for Information Science*, 27(3), 129-146. <https://doi.org/10.1002/asi.4630270302>
- Salmon, C. T., & Kline, F. G. (1983, May). *The spiral of silence ten years later: An examination and evaluation*. Annual Meeting of the International Communication Association, Dallas, TX, United States. <http://files.eric.ed.gov/fulltext/ED235528.pdf>
- Scheufle, D. A., & Moy, P. (2000). Twenty-five years of the spiral of silence: A conceptual review and empirical outlook. *International Journal of Public Opinion Research*, 12(1), 3-28. <https://doi.org/10.1093/ijpor/12.1.3>
- Scott, M. (2010). Problems in investigating keyness, or clearing the undergrowth and marking out trails. In M. Bondi & M. Scott (Eds.), *Keyness in texts* (pp. 43-58). John Benjamins Pub. <https://doi.org/10.1075/scl.41.04sco>
- Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior*, 7(3), 321-326. <https://doi.org/10.1089/1094931041291295>
- Sunstein, C. R. (2009). *Republic.com 2.0*. Princeton University Press. <https://a.co/d/g2Ge3oL>
- Taavitsainen, I., & Jucker, A. H. (2015). Twenty years of historical pragmatics: Origins, developments and changing thought styles. *Journal of Historical Pragmatics*, 16(1), 1-24. <https://doi.org/10.1075/jhp.16.1.01taa>
- Walther, J. B., & Jang, J. W. (2012). Communication processes in participatory websites. *Journal of Computer-Mediated Communication*, 18(1), 2-15. <https://doi.org/10.1111/j.1083-6101.2012.01592.x>
- Weninger, T., Johnston, T. J., & Glenski, M. (2015). Random voting effects in social-digital spaces: A case study of Reddit post submissions. *Proceedings of the 26th ACM conference on hypertext & social media*, 293-297. Association for Computing Machinery. <https://doi.org/10.1145/2700171.2791054>
- Wilson, C., & Irvine, K. N. (2013). Bottom-up communication: Identifying opportunities and limitations through an exploratory field-based evaluation. *Energy Efficiency*, 6, 91-104. <https://doi.org/10.1007/s12053-012-9161-y>
- Zerback, T., & Fawzi, N. (2017). Can online exemplars trigger a spiral of silence? Examining the effects of exemplar opinions on perceptions of public opinion and speaking out. *New Media & Society*, 19(7), 1034-1051. <https://doi.org/10.1177/1461444815625942>

Unveiling a New Approach to Opinion Climate Analysis in Diverse Information Environments

Yi-Lin Tsai*

ABSTRACT

This paper revisits Noelle-Neumann's Spiral of Silence theory, proposed in 1972, and identifies areas that need enhancement in the context of the contemporary information environment. By utilizing the Quasi-Statistical Organ from the theory, the study establishes a more suitable analytical framework for modern public opinion research. The following questions are explored. (1) How does the Quasi-Statistical Organ operate within the new media environment? (2) How can data be collected according to the logic of the Quasi-Statistical Organ? (3) Can observation methods designed based on the Quasi-Statistical Organ predict emerging popular topics?

The literature is divided into three sections: Theoretical Gaps Due to Changes in Media Environment, Social Cues in the New Media Environment, and New Research Approaches Combining New Media Opinion Climate and Corpus Methods.

First, the strand of Theoretical Gaps Due to Changes in Media Environment explores domestic and international studies on the application of the Spiral of Silence theory. It finds that changes in the information environment have enabled audiences to exhibit more diverse responses.

* Yi-Lin Tsai holds a Ph. D. from the College of Communication at National Chengchi University and will serve as an Assistant Professor in the Department of Mass Communication at Tamkang University starting February 2025. E-mail: elaine019691@gmail.com.

Whereas in the past, audiences could only speculate in interviews whether they would express their political choices to strangers on a train or participate in opinion polls, today, they can obtain social cues from various sources. Notably, the process of browsing social media allows audiences to piece together aggregate user representations (AURs), enabling them to imagine the majority opinion and decide whether to express their own views or subtly blend into AUR (Walther & Jang, 2012). Audiences now have more agency in the formation of opinion climate. More importantly, the behavior of detecting opinion climates through AUR helps researchers clarify the operation of the Quasi-Statistical Organ, a principle previously unexplained in the Spiral of Silence theory, in the contemporary media environment.

Second, the section Social Cues in the New Media Environment examines how, in different information environments, do proprietor content (PC), aggregate user representations (AURs), and user-generated content (UGC) constitute social cues from the audience's perspective.

Lastly, the part New Research Approaches Combining New Media Opinion Climate and Corpus Methods reviews research on public opinion to determine if there are methodological limitations, such as initiating studies from individual cases or topics, that could miss unanticipated research outcomes due to top-down research designs. This study argues that incorporating corpus methods, which include audience behavior of collecting opinion climates using AUR, into the data collection process and establishing a continuously updating research framework through linguistic probes align researchers' perspectives more closely with those of the audience.

This study selects the Gossip Board of PTT.cc, Taiwan's most popular bulletin board system, as the data source, using the aggregate user representations (AUR) accessible to users as the data collection standard. It collects data from January to June 2022 on PTT's Gossip Board, specifically

targeting posts that (1) have a net positive score (upvotes minus downvotes) greater than 100 and (2) have titles containing “Re: [新聞]” (Re: [News]), including any associated comments. The collected data comprise 651 posts and 319,876 comments. The data are organized and analyzed using the Chinese corpus analysis tool CORPRO. The criterion of selecting posts with a net positive score greater than 100 ensures that these posts are marked as “爆!” (explosive), making them more visible to users. Additionally, the requirement that the title contains “Re: [新聞]” indicates that the content is reposted news that has been brought back for discussion. This behavior is akin to audiences simultaneously looking towards the public domain and the community, receiving public news information, and then sharing it with others who should see it. This thought process closely aligns with the intuitive way audiences gauge others’ opinions on issues, as described in the Quasi-Statistical Organ. Moreover, the study builds upon Johansen’s (2017) method of using corpora to identify specific function words in the form₁-to-function-to-form₂ approach in probe studies, allowing researchers to continuously update probes when identifying emerging opinion climates.

The results of this paper establish a sustainable and topic-independent opinion analysis framework. First, the word frequency of a large number of comments is calculated. Second, replies to less popular posts are collected and subjected to keyness calculation to identify which words appear exclusively in popular posts. This set of words serves as probes capable of extracting the opinion climate. Third, articles containing these probes are searched over the following week, revealing the opinion climate currently under heated discussion. The research findings note that the analytical framework built using probes is indeed effective when applied to the six months of data collected herein.

The purpose of this study is to address gaps in the description of the

Quasi-Statistical Organ principle within the Spiral of Silence theory and to establish an opinion climate analysis framework that transcends time and topics, thus aligning closer to the perspectives of online community members. The primary research contributions include outlining the operational principles of the Quasi-Statistical Organ in the contemporary media environment and proposing a novel analytical framework for opinion climate analysis. This paper first addresses the deficiencies in the Spiral of Silence theory by outlining the operational principles of the Quasi-Statistical Organ in the current media environment and then presents a new, sustainable framework for opinion climate analysis that extends beyond predetermined research boundaries.

Based on this paper, future research can explore comparative studies on other platforms or serve as a reference for information literacy research. The limitations of this study include the inability to cover the impact of other social media interfaces and the inability to exclude any influences of some extreme cases on the overall data.

Keywords: aggregate user representation, opinion climate, quasi-statistical organ, corpus analysis

